



Gene 117, 249-254, 1992  
A;Title: Opioid-binding cell adhesion molecule (OBCAM)-related clones from a rat brain c  
A;Reference number: JCI1238; MUID:92347701; PMID:1339369  
A;Accession: JCI1238  
A;Molecule type: mRNA  
A;Residues: 1-338 <LIP>  
A;Cross-references: GB:M88709; NID:g203245; PIDN:AAA40858.1; PID:g203246  
A;Experimental source: brain  
C;Genetics:  
A;Gene: OBCAM  
C;Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 5.7%; Score 118.5; DB 2; Length 338;  
Best Local Similarity 25.2%; Pred. No. 0.22; Mismatches 107; Indels 71; Gaps 18;  
Matches 72; Conservative 36; Mismatches 107; Indels 71; Gaps 18;

Qy 5 LLTPEAVGSGSGNEVI-EGPQNATVLKGSQARENCTVSQGWKLIMWALSDMVVL----- 58  
Db 16 LLFIP-GVPRSGDATEPKAMDNTVVRQGESATLRCITDDRETVAV-LNRSTILYAGND 73

Qy 59 --SVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVEPDSGNI RCSLQ-----NSRLH 110  
Db 74 KWSIDPRVILVN---TPTQY-----SIMIQNVVDYDEGPTCSVQTDNHPKTSRVH 122

Qy 111 GSAVLTQVMGELFIPVNLVVAENEPEVTCL----PSHWTRLPDISWELGLLVSHSSY 166  
Db 123 ----LIVQPPQIMNISDIITVNEISSVTLLCLAIGRPE-----PTVTWR-----HLSV 167

Qy 167 Y----FVPEPSDLOSAYSIIALTPQSNGLTCVATWKS LKARKSATVNLTV-----IRCP 217  
Db 168 KEGQGFVSEDEYLE----ISDIKRQSGEYEC SAL-NDVAAPDVVRKVKITVNYPPYISKA 222

Qy 218 QDTGGGINIPQVL-----SSLPSLGFSLPTWCK-----VGLGLAGTML 255  
Db 223 KNTGVSVGQKILSCEASAVPMAEFQ---WPKEDTRLATGLDGVRI 265

Search completed: November 13, 2003, 03:34:29  
Job time : 16.3865 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 23:23:44 ; Search time 9.36568 Seconds  
(without alignments)  
1938.172 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077

Sequence: 1 MERHLTVPEAVGSGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	7.3	404	1	RAGE_HUMAN
2	139.5	6.7	344	1	NTRI_RAT
3	133.5	6.4	344	1	NTRI_HUMAN
4	133.5	6.4	344	1	NTRI_MOUSE
5	133	6.4	337	1	G55A_CHICK
6	131	6.3	1447	1	DCC_MOUSE
7	130	6.3	345	1	OPCM_HUMAN
8	129	6.2	345	1	OPCM_BOVIN
9	129	6.2	353	1	CEPU_CHICK
10	129	6.2	6632	1	UN89_CABEL
11	128	6.2	345	1	OPCM_RAT
12	128	6.2	1447	1	DCC_HUMAN
13	128	6.2	1914	1	KMLS_HUMAN
14	125	6.0	1070	1	PTK7_HUMAN
15	122.5	5.9	4391	1	PGBM_HUMAN
16	118.5	5.7	620	1	SMP_COTUA
17	117.5	5.7	1443	1	NEOI_CHICK
18	114.5	5.5	1461	1	NEOI_HUMAN
19	114	5.5	569	1	SILF_MOUSE
20	113.5	5.5	416	1	RAGE_BOVIN
21	113.5	5.5	1377	1	NEOI_RAT
22	111	5.3	365	1	CKAR_HUMAN
23	110.5	5.3	524	1	BUTY_MOUSE
24	110.5	5.3	764	1	ICCR_DROME
25	109.5	5.3	249	1	CSP_DROME
26	109.5	5.3	890	1	TYO3_HUMAN
27	106	5.1	1092	1	NCA2_XENLA
28	105.5	5.1	862	1	CD22_MOUSE
29	105.5	5.1	3707	1	PGBM_MOUSE
30	105	5.1	319	1	A33_HUMAN
31	104.5	5.0	333	1	AMAL_DROME
32	104.5	5.0	879	1	FFRP_RAT
33	104	5.0	879	1	FFRP_MOUSE

#### ALIGNMENTS

##### RESULT 1

RAGE\_HUMAN STANDARD; PRT; 404 AA.  
AC Q15109; Q15279; Q9H2X7; Q9Y3R3;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Advanced glycosylation end product-specific receptor precursor  
DE (Receptor for advanced glycosylation end products).  
GN AGER OR RAGE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Lung;  
RX MEDLINE=92340547; PubMed=1378843;  
RA Neepser M., Schmidt A.M., Brett J., Van S.D., Wang F., Pan Y.C.,  
RA Elliston K., Stern D., Shaw A.;  
RT "Cloning and expression of a cell surface receptor for advanced  
RT glycosylation end products of proteins.";  
RL J. Biol. Chem. 267:14998-15004(1992).  
[2]  
SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=95137587; PubMed=7835890;  
RA Sugaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,  
RA Inoko H., Ikemura T.;  
RT "Three genes in the human MHC class III region near the junction with  
RT the class II: gene for receptor of advanced glycosylation end  
RT products, P8X2 homeobox gene and a notch homolog, human counterpart  
RT of mouse mammary tumor gene int-3.";  
RL Genomics 23:408-419(1994).  
[3]  
SEQUENCE FROM N.A. (ISOFORM 1).  
RA Rowen L., Danks C., Baskin D., Faust J., Loretz C., Ahearn M.E.,  
RA Banta A., Spies T., Hood L.;  
RT "Sequence determination of 300 kilobases of the human class III MHC  
RT locus.";  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
[4]  
SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.  
RA Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,  
RA Yamamoto H.;  
RT "Molecular heterogeneity of the receptor for advanced glycation  
RT endproducts.";  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE FROM N.A. (ISOFORM 2).  
RA Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,  
RA Schuler A., Huber G.;  
RT "cDNA cloning of a novel secreted isoform of the human receptor for  
RT advanced glycation end products (RAGE) and characterization of cells  
RT co-expressing cell-surface scavenger receptors and Swedish mutant  
RT amyloid precursor protein.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

Q98919 gallus gall  
P08921 rattus norv  
Q92038 rattus norv  
P97792 mus musculu  
O43699 homo sapien  
Q14162 homo sapien  
P11799 gallus gall  
Q06561 caenorhabdi  
P34082 drosophila  
P14781 gallus gall  
O60500 homo sapien  
P55146 rattus norv

34 103.5 5.0 338 1 LAMP\_CHICK  
35 102.5 4.9 344 1 CD2\_RAT  
36 102.5 4.9 348 1 KILQ\_RAT  
37 102.5 4.9 365 1 CXAR\_MOUSE  
38 100.5 4.8 442 1 SIL6\_HUMAN  
39 100.5 4.8 830 1 SREC\_HUMAN  
40 100.5 4.8 1906 1 KMLS\_CHICK  
41 100.5 4.8 3375 1 UN52\_CABEL  
42 100 4.8 873 1 PAS2\_DROME  
43 100 4.8 1010 1 CONT\_CHICK  
44 100 4.8 1241 1 NPHN\_HUMAN  
45 99.5 4.8 880 1 TYO3\_RAT



DE Neurotrophin precursor (GP65).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 217-229.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=95198094; PubMed=7891157;  
 RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,  
 RA Salzer J.L.;  
 RT Cloning of neurotrophin defines a new subfamily of differentially  
 RT expressed neural cell adhesion molecules.";  
 RL J. Neurosci. 15:2141-2156(1995).  
 CC -1- FUNCTION: Neural cell adhesion molecule.  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN SEVERAL  
 CC SUBPLATE, AND LOWER CORTICAL LAMINAE IN THE FOREBRAIN AND IN THE  
 CC PONTINE NUCLEUS, CEREBELLAR GRANULE CELLS, AND PURKINJE CELLS IN  
 CC THE HINDRAIN.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U16845; AAA67445.1; -.  
 DR PIR; I56551.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00408; IGC2; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 3.  
 DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
 KW Repeat; Signal  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 321 NEUROTROPHIN.  
 FT PROPEP 322 344 REMOVED IN MATURE FORM (POTENTIAL).  
 FT DOMAIN 39 126 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 136 218 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 222 309 IG-LIKE C2-TYPE 3.  
 FT DISULFID 57 115 POTENTIAL.  
 FT DISULFID 157 201 POTENTIAL.  
 FT DISULFID 243 295 POTENTIAL.  
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT LIPID 321 321 GPI-ANCHOR (POTENTIAL).  
 SQ SEQUENCE 344 AA; 37998 MW; CBB39BE53B3B224 CRC64;

Query Match 6.7%; Score 139.5; DB 1; Length 344;  
 Best Local Similarity 24.9%; Pred. No. 0.00073;  
 Matches 66; Conservative 40; Mismatches 110; Indels 49; Gaps 13;  
 QY 3 RHLLTPEAVGSGSGNEVI-EGPQNAVTLKGSQARENCTVSGQWKLIMWALSMDVLS-- 59  
 DB 20 RLFLVPTGVPSGGATFFPAMDNVTVRGESATURCTIDNRVTRVAVLNRSTILYAGN 79  
 QY 60 ---VRPMEPIITNDRFTSQYDQGGNFTSEMIIHNVPSDSGNIRCSLQ-----NSRL 109

Db 80 DKWCLDRVVLN-----SIEIQNVVDYDEGPTCSVQDTHPKTSRV 128  
 QY 110 HGSAYLTQVMGELFSPVNLVAENPCBVTCLPSSHWTRLPDLSWELGLLVSHSSYFV 169  
 Db 129 H-----LIVQSPKIVEISSDISINEGNNISLTCTIGRPE-PTVTWR---HISPKAVGFV 180  
 QY 170 PEPSDLOSAYSIIALTQPSNGTLTCVATWKSLSKARKSATVNLTVIRCP-----QDTGGGI 224  
 Db 181 SEDEYLE---IQGITREQSGEYCSAS-NDVAAPVVRVNVVTPYPISEAKGTGVPV 235  
 QY 225 NIQVVL-----SSLPSLGSLPTWCK 245  
 Db 236 GQKGTLCQCEASVPSAEFQ---WFK 257  
 RESULT 3  
 NTRI HUMAN  
 ID NTRI\_HUMAN STANDARD; PRT; 344 AA.  
 AC Q9P121;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Neurotrophin precursor (hnt).  
 GN NT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Li G., Jin J., Tan X., Hu S., Yuan J., Qiang B.;  
 RT Cloning and identification of human neurotrophin full length cDNA.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Neural cell adhesion molecule.  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF126426; AAF37591.1; -.  
 DR GO; GO:0008038; P:neuronal cell recognition; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00408; IGC2; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 3.  
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
 KW Repeat; Signal  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 321 NEUROTROPHIN.  
 FT PROPEP 322 344 REMOVED IN MATURE FORM (POTENTIAL).  
 FT DOMAIN 39 126 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 136 218 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 222 309 IG-LIKE C2-TYPE 3.  
 FT DISULFID 57 115 POTENTIAL.  
 FT DISULFID 157 201 POTENTIAL.  
 FT DISULFID 243 295 POTENTIAL.  
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 321 321 N-LINKED (GLCNAC...) (POTENTIAL).
FT LIPID 321 321 GPI-ANCHOR (POTENTIAL).
SQ SEQUENCE 344 AA; 37971 MW; DA4D12C295ABBE3A CRC64;

Query Match 6.4%; Score 133.5; DB 1; Length 344;
Best Local Similarity 24.5%; Pred. No. 0.0022;
Matches 65; Conservative 40; Mismatches 111; Indels 49; Gaps 13;

QY 3 RHLLVTPAVGSGSGNEVI-EGPQNAVTLKGSQARFNCVTSGQWKLIMWALSDMVVLS-- 59
DB 20 RLLFLVPTGVPVRSQDGFPPKAMDNVTVRQGESATLRCITIDNRVTRVAVLNRSTILYAGN 79
QY 60 ---VRPMEPIITNDRFTSQRYDQGNFTSEMIHNVPDSGNIRCSLQ-----NSRL 109
DB 80 DKWCLDPRVLLSN---TQTQV-----STEIQNVVDVDEGPTCSQVTDNHPKTSRV 128
QY 110 HGSAYLTVOVMGELFIPSVNLVVAENPECEVTLPSHWTRLPDISWELGLLVSHSYVYFV 169
DB 129 H----LIVQVSPKIVEISSDISINEGNNISLTCIATGRPE-PTVTWR---HISPKAVGFV 180
QY 170 PPSPLQSAVSTLALTPQSNGLTLCVATWKSLSKARKSATVNLTVIRCP-----QDTGGGI 224
DB 181 SEDEYLE----IQGITREQSGEYECAS--NDVAAPVVRVVKTVVNPYPPISEAKGTGVPV 235
QY 225 NIPGVL----SSLSPLGFSPLPTWGK 245
DB 236 GQKGTLOCEASAVPSAEFQ---WYK 257

RESULT 4
NTRI MOUSE STANDARD; PRT; 344 AA.
AC Q99PJ0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurotrimin precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RA Kim T.H., Choi S.C., Kim J., Jeon J.W., Kim K.D., Lee S.H.;
RT "Cloning and expression of mouse neurotrimin gene in the developing
RT nervous system.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Walek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Neural cell adhesion molecule.
```

```
CC CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC CC SUBFAMILY.
CC CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AF282980; AAK00276.1; --
CC CC EMBL; BC023307; AAK23307.1; --
CC CC InterPro; IPR007110; Ig-like.
CC CC InterPro; IPR003599; Ig.
CC CC InterPro; IPR003598; Ig_c2.
CC CC InterPro; IPR003006; Ig_MHC.
CC CC Pfam; PF00047; Ig; 3.
CC CC SMART; SM00408; IG; 3.
CC CC SMART; SM00408; IGC2; 3.
CC CC PROSITE; PS50835; IG_LIKE; 3.
CC CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.
FT SIGNAL 1 31
FT CHAIN 32 321
FT PROPEP 322 344
FT DOMAIN 39 126
FT DOMAIN 136 218
FT DOMAIN 222 309
FT DISULFID 57 115
FT DISULFID 157 201
FT DISULFID 243 295
FT CARBOHYD 44 44
FT CARBOHYD 70 70
FT CARBOHYD 152 152
FT CARBOHYD 284 284
FT CARBOHYD 292 292
FT CARBOHYD 305 305
FT CARBOHYD 321 321
FT LIPID 321 321
FT CONFLICT 75 75
FT CONFLICT 92 92
FT CONFLICT 119 119
FT CONFLICT 187 187
FT CONFLICT 213 213
FT CONFLICT 225 225
SQ SEQUENCE 344 AA; 37984 MW; C885BBA52C148554 CRC64;

Query Match 6.4%; Score 133.5; DB 1; Length 344;
Best Local Similarity 24.5%; Pred. No. 0.0022;
Matches 65; Conservative 40; Mismatches 111; Indels 49; Gaps 13;

QY 3 RHLLVTPAVGSGSGNEVI-EGPQNAVTLKGSQARFNCVTSGQWKLIMWALSDMVVLS-- 59
DB 20 RLLFLVPTGVPVRSQDGFPPKAMDNVTVRQGESATLRCITIDNRVTRVAVLNRSTILYAGN 79
QY 60 ---VRPMEPIITNDRFTSQRYDQGNFTSEMIHNVPDSGNIRCSLQ-----NSRL 109
DB 80 DKWCLDPRVLLSN---TQTQV-----STEIQNVVDVDEGPTCSQVTDNHPKTSRV 128
QY 110 HGSAYLTVOVMGELFIPSVNLVVAENPECEVTLPSHWTRLPDISWELGLLVSHSYVYFV 169
DB 129 H----LIVQVSPKIVEISSDISINEGNNISLTCIATGRPE-PTVTWR---HISPKAVGFV 180
QY 170 PPSPLQSAVSTLALTPQSNGLTLCVATWKSLSKARKSATVNLTVIRCP-----QDTGGGI 224
DB 181 SEDEYLE----IQGITREQSGEYECAS--NDVAAPVVRVVKTVVNPYPPISEAKGTGVPV 235
QY 225 NIPGVL----SSLSPLGFSPLPTWGK 245
DB 236 GQKGTLOCEASAVPSAEFQ---WYK 257
```

```

Best Local Similarity 25.9%; Pred. No. 0.0023;
Matches 71; Conservative 36; Mismatches 107; Indels 60; Gaps 16;

QY 5 LLTPVAVGGSGNEVI-EGPQNAIVLKGSQARFNCTVSGQKLIIMWALSDMVL----- 58
DB 16 LLFIP-GVPRSGDATFPKAMDNVTVRQGSATURCTVDDRRVRVAM-LNRSTLIYAGND 73
QY 59 --SVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVFSDSGNIRCSLQ-----NSRLH 110
DB 74 KWSIDNRWILSN--TKTQY-----SIKIHNVDYDGPVTCSTQDNHPKTSRVH 122
QY 111 GSAYLTQVMGELFIPSNVLVAENPEVTCVCL-----PSHWTRLPDISWELGLLVSHSSY 166
DB 123 ----LIVQVPPQIVNISSDITVNEGSSVTLMCLAFGRPE-----FTVTWR---HLSGRGQ 170
QY 167 YFVPEPSDLSAVSILALTPQSNGLTLCVATKWKSKARKSATVNLTV-----IRCPQDTG 221
DB 171 GFVSEDEYLE-----ITGITREQSGEYECASV-NDVAVPDVVRKVKVTVNPYPPIISNAKNTG 225
QY 222 GGINIPGVL-----SSLSLGFSLFTWCKVGLGLA 251
DB 226 ASVGQKGILOCEASAVPVAEFQ-----WFKEDTRLA 256

RESULT 6
DCC MOUSE STANDARD; PRT; 1447 AA.
AC P70211;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tumor suppressor protein DCC precursor.
GN DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RC STRAIN=BALB/c; TISSUE=Brain;
RC MEDLINE=96112625; PubMed=8570174;
RA Cooper H.M., Ames P., Britto J., Gad J., Wilks A.F.;
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RT gene (mDCC) and its expression in the developing mouse embryo."
RL Oncogene 11:2243-2254(1995).
RN [2]
RP REVISIONS.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Evt=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=P70211-1; Sequence=Displayed;
CC Note=Isoform B is produced by alternative initiation at Met-85
CC of isoform A;
CC Name=C;
CC IsoId=P70211-2; Sequence=VSP_002501;
CC Evt=Alternative initiation;
CC Comment=2 isoforms, A (shown here) and B, are produced by
CC alternative initiation at Met-1 and Met-85;
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.
CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION
CC AND REMAIN AT THIS LEVEL IN THE ADULT.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.

```



```
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 322 322 GPI-ANCHOR (POTENTIAL).
SQ SEQUENCE 345 AA; 38007 MW; E7AD17BEA1AA3FF4 CRC64;

Query Match
Best Local Similarity 6.3%; Score 130; DB 1; Length 345;
Matches 73; Conservative 34; Mismatches 108; Indels 70; Gaps 17;

QY 3 RLLTVPVAVGSGGNEVI-EGPONATVVKGSQARFNCVTSQGWKLINWALSDMVVL--- 58
Db 20 RLLFLVPTGVPRSGDATFPKMDNVTVRQGESATLRCITDIDRVTRVAV-LNRSTILYAG 78
QY 59 ----SVRPMPEIITNDRTSQRVQGGNFTSEMIHNHVPDSGNIRCSLQ-----NSR 108
Db 79 NDKWSIDPRVILVN---TPTQY-----SIMQNVVDVDEGFTCSVQTDNHPKTSR 127
QY 109 LHGSAYLTQVMGELFIPSNLVVAENEPCEVTCL-----PSHWTRLPDISWELGLLVSHS 164
Db 128 VH-----LIVQVPPQIMNISSDITVNEGSSVTLCLAIGRPE-----PTVTWR-----HL 172
QY 165 SY-----FVPEPSDLSQSAVSLALTPQSGNGLTCTVATWKSILKARKSATVNLTV-----IR 215
Db 173 SVKEGQGFVSEDEYLE-----ISDIKRDQSGEYECAL-NDVAAPDVRVKVITVNPYPVIS 227
QY 216 CPQDTGGGINIPGVL-----SSLPSLGFSLPTWKGVLGLA 251
Db 228 KAKNTGVSQKGLSCSAVPMAEFQ---WFKEDTFLA 264

RESULT 8
ID OPM BOVIN STANDARD; PRT; 345 AA.
AC P11834;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OBCAM)
GN OPM OR OBCAM OR OCAM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=89251576; PubMed=2721489;
RA Schofield P.R., McFarland K.C., Hayflick J.S., Wilcox J.N.,
RA Cho T.M., Roy S., Lee N.M., Loh H.H., Seeburg P.H.;
RT "Molecular characterization of a new immunoglobulin superfamily
RL protein with potential roles in opioid binding and cell contact.";
RL EMBO J. 8:489-495 (1989).
CC -!- FUNCTION: Binds opioids in the presence of acidic lipids; probably
CC involved in cell contact.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X12672; CAA31192.1; --
CC PIR; S03199; S03199.
DR
```

```
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; IG2; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 322
FT PROPEP 323 345
FT DOMAIN 39 126
FT DOMAIN 136 219
FT DOMAIN 223 310
FT DISULFID 57 115
FT DISULFID 157 202
FT DISULFID 244 296
FT CARBOHYD 44 44
FT CARBOHYD 70 70
FT CARBOHYD 140 140
FT CARBOHYD 285 285
FT CARBOHYD 293 293
FT CARBOHYD 306 306
FT LIPID 322 322
SQ SEQUENCE 345 AA; 37914 MW; DIECC8D9E7D8CB19 CRC64;

Query Match
Best Local Similarity 6.2%; Score 129; DB 1; Length 345;
Matches 72; Conservative 33; Mismatches 109; Indels 66; Gaps 16;

QY 3 RLLTVPVAVGSGGNEVI-EGPONATVVKGSQARFNCVTSQGWKLINWALSDMVVL--- 58
Db 20 RLLFLVPTGVPRSGDATFPKMDNVTVRQGESATLRCITDIDRVTRVAV-LNRSTILYAG 78
QY 59 ----SVRPMPEIITNDRTSQRVQGGNFTSEMIHNHVPDSGNIRCSLQ-----NSR 108
Db 79 NDKWSIDPRVILVN---TPTQY-----SIMQNVVDVDEGFTCSVQTDNHPKTSR 127
QY 109 LHGSAYLTQVMGELFIPSNLVVAENEPCEVTCL-----PSHWTRLPDISWELGLLVSHS 164
Db 128 VH-----LIVQVPPQIMNISSDITVNEGSSVTLCLAIGRPE-----PTVTWR-----HL 172
QY 165 SY-----FVPEPSDLSQSAVSLALTPQSGNGLTCTVATWKSILKARKSATVNLTV-----IR 215
Db 173 SVKEGQGFVSEDEYLE-----ISDIKRDQSGEYECAL-NDVAAPDVRVKVITVNPYPVIS 227
QY 216 CPQDTGGGINIPGVL-----SSLPSLGFSLPTWKGVLGLA 251
Db 228 KAKNTGVSQKGLSCSAVPMAEFQ---WFKEDTFLA 264

RESULT 9
ID CEPU CHICK STANDARD; PRT; 353 AA.
AC Q90773;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CEPU-1 protein precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=96370549; PubMed=8774445;
RA Spaltmann F., Bruemendorf T.;
RT "CEPU-1, a novel immunoglobulin superfamily molecule, is expressed by
RL developing cerebellar Purkinje cells.";
RL J. Neurosci. 16:1770-1779 (1996).
```

CC -!- FUNCTION: IT MAY BE A CELLULAR ADDRESS MOLECULE SPECIFIC TO  
 CC PURKINJE CELLS. IT MAY REPRESENT A RECEPTOR OR A SUBUNIT OF A  
 CC RECEPTOR COMPLEX.  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=Minor;  
 CC IsoId=Q90773-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Major;  
 CC IsoId=Q90773-2; Sequence=VSP\_002607;  
 CC -!- TISSUE SPECIFICITY: FOUND ON THE DENDRITES, SOMATA AND AXONS OF  
 CC DEVELOPING PURKINJE CELLS. UNDETECTABLE ON OTHER NEURONS LIKE  
 CC GOLGI OR GRANULE CELLS.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED BY DEVELOPING CEREBELLAR PURKINJE  
 CC CELLS. EXPRESSION COINCIDES WITH THE GROWTH OF THE DENDRITIC TREE,  
 CC AFTER PURKINJE CELLS HAVE FINISHED THEIR MIGRATION FROM THE  
 CC VENTRICULAR ZONE (FROM E15 UNTIL E21). EXPRESSED IN THE ADULT.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; Z72497; CAA96578.1; --  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_C2.  
 DR Pfam; PF00047; Ig\_3.  
 DR SMART; SM00408; IGG2; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 3.  
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
 KW Repeat; Signal; Alternative splicing.  
 FT SIGNAL 28  
 FT CHAIN 29 330 CEPU-1 PROTEIN.  
 FT PROPEP 331 353 REMOVED IN MATURE FORM (POTENTIAL).  
 FT DOMAIN 37 124 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 134 216 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 220 314 IG-LIKE C2-TYPE 3.  
 FT DISULFID 55 113 POTENTIAL.  
 FT DISULFID 155 199 POTENTIAL.  
 FT DISULFID 241 293 POTENTIAL.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT LIPID 330 330 GPI-ANCHOR (POTENTIAL).  
 FT VARSPLIC 310 320 Missing (in isoform 2).  
 FT SEQUENCE 353 AA; 38736 MW; 255048591EBB46 CRC64;  
 Query Match 6.2%; Score 129; DB 1; Length 353;  
 Best Local Similarity 25.0%; Pred. No. 0.0051;  
 Matches 66; Conservative 38; Mismatches 92; Indels 68; Gaps 14;  
 QY 7 TVPEAVSGSGNEVIEGPNATVLKGSQARNCTVSO-----GW---KLINWALSDMVVL 58  
 Db 35 TFPKAM-----DNTVTRQGESATLRCSVDNRVTRVAVLNRSLSIYAQNDKWL 82  
 QY 59 SVRPWEPIITNDRTSORYDQGNFTSEMITHNVPDSGNIRCSLO-----NSRLHGS 112  
 Db 83 DPRVVLANTKIQSIQ-----IHDVDVDEGPTCSVDNHPKTSRVH-- 127  
 QY 113 AYLTVQVNGELFIPSVNLVAENPECVETCLPWSHTRLPD--ISWELGLLVSHSSYFVP 170  
 Db 128 --LIVQVSPKITETSSDISINEGGVSLTICIA---TGRPDPTITWR---HISPKAVGFIS 179

QY 171 EPSDLOSAVSILALTQPSNGTLTCVATWKSARKSATVNLTV-----IRCPDQTGGGIN 225  
 Db 180 EDEYLE---ITGITREQSGEYECSSAS-NDVAAPVVQVRKVTVNYPYISDAKSTGVPVG 234  
 QY 226 IPGVL-----SSPSLGFSLPTWCK 245  
 Db 235 QKGILMCEASAVPSADFQ---WYK 255  
 RESULT 10  
 ID UN89 CAEEL STANDARD; PRT; 6632 AA.  
 AC O01761; Q17362;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).  
 GN UNC-89 OR C09D1.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=96180278; PubMed=8603916;  
 RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;  
 RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line  
 RT assembly, encodes a giant modular protein composed of Ig and signal  
 RT transduction domains.";  
 RL J. Cell Biol. 132:835-848 (1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Du Z., Le T.T., Wilson R.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Structural component of the muscle M-line. Myofilament  
 CC lattice assembly begins with positional cues laid down in the  
 CC basement membrane and muscle cell membrane. UNC-89 responds to  
 CC these signals, localizes, and then participates in assembling an  
 CC M-line.  
 CC -!- TISSUE SPECIFICITY: Localizes to the middle of A-bands.  
 CC -!- SIMILARITY: Contains 1 DSL-homology (DH) domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type III domain.  
 CC -!- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.  
 CC -!- SIMILARITY: Contains 1 PH domain.  
 CC -!- SIMILARITY: Contains 5 RCSD domains.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; U33058; AAB00542.1; --  
 DR EMBL; AF003131; AAB54132.2; --  
 DR PDB; 1FHO; 20-DEC-00.  
 DR WormPep; C09D1.1; CE30426  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_C2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR007850; RCSD.  
 DR InterPro; IPR000219; RhoGEF.



RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=92347701; PubMed=1339369;  
RA Lippman D.A., Lee N.M., Loh H.H.;  
RT "Opioid-binding cell adhesion molecule (OBCAM)-related clones from a  
RT rat brain cDNA library."  
RL Gene 117:249-254(1992).  
RN [2]  
RP SEQUENCE OF 195-214, AND GPI-ANCHOR.  
RX MEDLINE=95198094; PubMed=7891157;  
RA Struyk A.F., Canoli P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,  
RA Salzer J.L.;  
RT "Cloning of neurotrophin defines a new subfamily of differentially  
RT expressed neural cell adhesion molecules."  
RL J. Neurosci. 15:2141-2156(1995).  
CC -!- FUNCTION: Binds opioids in the presence of acidic lipids; probably  
CC involved in cell contact.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1;  
CC IsoId=P32736-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=P32736-2; Sequence=VSP\_002612;  
CC Note=No experimental confirmation available;  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M88710; AAA40859.1; -;  
DR EMBL; M88711; AAA40860.1; -;  
DR EMBL; M88709; AAA40858.1; -;  
DR PIR; JCI238; JCI238.  
DR PIR; JCI239; JCI239.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig C2.  
DR Pfam; PF000047; ig; 3.  
DR SMART; SM00408; IGG2; 2.  
DR PROSITE; PS50835; IG-LIKE; 3.  
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
KW Repeat; Signal; Alternative splicing.  
FT SIGNAL 1 27  
FT CHAIN 28 322  
FT  
FT OPTOID BINDING PROTEIN/CELL ADHESION  
FT MOLECULE.  
FT  
FT REMOVED IN MATURE FORM (POTENTIAL).  
FT  
FT IG-LIKE C2-TYPE 1.  
FT  
FT IG-LIKE C2-TYPE 2.  
FT  
FT IG-LIKE C2-TYPE 3.  
FT  
FT DISULFID 57 115  
FT  
FT DISULFID 157 202  
FT  
FT DISULFID 244 296  
FT  
FT CARBOHYD 44 44  
FT  
FT CARBOHYD 70 70  
FT  
FT CARBOHYD 140 140  
FT  
FT CARBOHYD 285 285  
FT  
FT CARBOHYD 293 293  
FT  
FT CARBOHYD 306 306  
FT  
FT CARBOHYD 322 322  
FT  
FT LIPID  
FT VARSPLIC 1 27  
MGVCGYFLPFWKCLVVVSLFLVPT -> MYHPAYWIVF  
SATALLFIP (in isoform 2).  
/FTId=VSP\_002612.  
SEQUENCE 345 AA; 38067 MW; A3181B0753F9658E CRC64;

Query Match 6.2%; Score 128; DB 1; Length 345;  
Best Local Similarity 25.3%; Pred. No. 0.006;  
Matches 73; Conservative 35; Mismatches 110; Indels 70; Gaps 17;  
QY 3 RHLTVPEAVGSGGNEVI-EGPQNAVTLKSGQARFNCTVSGWKLMALMSDMVVL--- 58  
DB RLLFLVPTGTPVRSRGDAPFKAMDNVTVRQGESATRLCIDRVTVRVAV-LNRSTILYAG 78  
QY 59 -----SVRPMPEIITNDRFTSQRYDQGNFTSEMIHNVEPDSGNTSCSLQ-----NSR 108  
DB 79 NDKWSIDPRVILVN---TPTQY-----SIMIQNVVDYDEGPTCSQTDNHPKTSR 127  
QY 109 LHGSALVTQVWGELEFIPSNLVVAENPECEVTCL-----PSHWTRLPDISWELGLLVSHS 164  
DB 128 VH-----LVQVPPQIMNISSDITVNEISSVLLCLLAIGRPE-----PTVTWR-----HL 172  
QY 165 SYV-----FVPEPSDLOSASVILALTPQSNGLTLCVATWKSLLKARKSATVNLTV-----IR 215  
DB 173 SVKEGGQFVSEDEYLE-----ISDKRDQSGEVECSAL-NDVAAPDVRKVKITVNYPPYIS 227  
QY 216 CPQDTGGGINIPQVL-----SSLPSLGFSLPTWCK-----VGLGLAGTML 255  
DB 228 KAKNTGVSQGGILSCBASAVPMAEFQ---WFKEDTRLATGLDGVRI 272  
RESULT 12  
DCC\_HUMAN  
ID DCC\_HUMAN STANDARD; PRT; 1447 AA.  
AC P43146;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tumor suppressor protein DCC precursor (Colorectal cancer suppressor).  
GN DCC.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95011532; PubMed=7926722;  
RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,  
RA Vogelstein B.;  
RT "The DCC gene product in cellular differentiation and colorectal  
RT tumorigenesis."  
RL Genes Dev. 8:1174-1183(1994).  
RN [2]  
RP SEQUENCE OF 1-750 FROM N.A.  
RX MEDLINE=90100559; PubMed=2294591;  
RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,  
RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,  
RA Vogelstein B.;  
RT "Identification of a chromosome 18q gene that is altered in  
RT colorectal cancers."  
RL Science 247:49-56(1990).  
RN [3]  
RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBELD EXONS).  
RX MEDLINE=91121517; PubMed=1991322;  
RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,  
RA Oliner J.D., Kinzler K.W., Vogelstein B.;  
RT "Scrambled exons."  
RL Cell 64:607-613(1991).  
RN [4]  
RP GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.  
RX MEDLINE=94245241; PubMed=8188295;  
RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,  
RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;  
RT "The DCC gene: structural analysis and mutations in colorectal  
RT carcinomas."  
RL Genomics 19:525-531(1994).  
RN [5]  
RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.

RX MEDLINE=94243823; PubMed=8187090;  
 RA Miyake S., Nagai K., Yoshino K., Oto M., Endo M., Yuasa Y.;  
 RT "Point mutations and allelic deletion of tumor suppressor gene DCC in  
 human esophageal squamous cell carcinomas and their relation to  
 metastasis";  
 RL Cancer Res. 54:3007-3010(1994).  
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL  
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.  
 CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO  
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC  
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR  
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS  
 CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 6 fibronectin type III domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL; X76132; CAA53735.1; -;  
 DR EMBL; M32292; AAA35751.1; -;  
 DR EMBL; M32286; AAA52174.1; -;  
 DR EMBL; M32288; AAA52175.1; ALT\_SEQ.  
 DR EMBL; M32290; AAA52176.1; -;  
 DR EMBL; M63696; AAA52177.1; -;  
 DR EMBL; M63700; AAA52178.1; -;  
 DR EMBL; M63702; AAA52179.1; -;  
 DR EMBL; M63718; AAA52180.1; -;  
 DR EMBL; M63698; AAA52181.1; -;  
 DR PIR; A54100; A54100.  
 DR HSSP; P56276; 1TLK.  
 DR Genew; HGNC:2701; DCC.  
 DR MIM; 120470; -;  
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO; GO:0007409; P:axogenesis; TAS.  
 DR GO; GO:0006917; P:induction of apoptosis; TAS.  
 DR GO; GO:0007048; P:oncogenesis; TAS.  
 DR InterPro; IPR003961; FN III.  
 DR InterPro; IPR003962; FNIII subd.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_C2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00041; fn3; 6.  
 DR Pfam; PF00047; ig; 4.  
 DR PRINTS; PR00014; FNTYPRIII.  
 DR SMART; SM00060; FN3; 6.  
 DR SMART; SM00408; IGC2; 3.  
 DR PROSITE; PS50835; IG-LIKE; 4.  
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;  
 KW Repeat; Anti-oncogene; Disease mutation; Polymorphism.  
 FT SIGNAL 1 25  
 FT CHAIN 26 1447  
 FT DOMAIN 26 1097  
 FT TRANSEM 1098 1122  
 FT DOMAIN 1123 1447  
 FT DOMAIN 26 135  
 FT DOMAIN 139 229  
 FT DOMAIN 234 326  
 FT DOMAIN 331 416  
 FT DOMAIN 426 522  
 FT DOMAIN 525 618  
 FT DOMAIN 619 716  
 FT DOMAIN 722 816  
 FT DOMAIN 840 940

FT DOMAIN 941 1042  
 FT DISULFID 61 117  
 FT BY SIMILARITY.  
 FT DISULFID 161 212  
 FT BY SIMILARITY.  
 FT DISULFID 352 400  
 FT BY SIMILARITY.  
 FT CARBOHYD 94 94  
 FT CARBOHYD 299 299  
 FT CARBOHYD 318 318  
 FT CARBOHYD 478 478  
 FT CARBOHYD 628 628  
 FT CARBOHYD 702 702  
 FT VARIANT 168 168  
 FT VARIANT 201 201  
 FT VARIANT 1375 1375  
 FT CONFLICT 138 138  
 FT CONFLICT 233 329  
 FT CONFLICT 421 421  
 SQ SEQUENCE 1447 AA; 158456 MW; 4A8612766ED0471F CRC64;  
 Query Match 6.2%; Score 128; DB 1; Length 1447;  
 Best Local Similarity 24.8%; Pred. No. 0.036; Indels 50; Gaps 10;  
 Matches 60; Conservative 33; Mismatches 99;  
 Qy 14 SSGGNEY-----IEGPNATVLKSGQARFNCTVSQGW--KLIMWALSDM 55  
 Db 220 SRTGNEAEVRLSDPGLHRLQYFLQPSNVVAIEGKDAVLECCVS-GYPFPTWLRGEE 278  
 Qy 56 VVLSVRPMEPIITNDRTSORYDQGNFTSEMTIHNVPSDSGNIRC--SLQNSRLHGS 113  
 Db 279 VI-----QLRSKKYSLLGG--SNLLISNVTDSDSGMYTCVTVTKNENISASA 323  
 Qy 114 YLTVOVMGELFIPSNVLVAENEPCEVTCLPSPHWRPLDISW-ELGLLVSHSSYFVPEP 172  
 Db 324 ELTVLPVPFWFLNHPNSNLYAYESMDIEFECTVS-GKPTVTNMKNGDVIFSDIPQIVGG 382  
 Qy 173 SDLOSASVILALTPQSGNLTTCVATKSLKARKSATVNLTVIRCPQDTGGGINPGVLSS 232  
 Db 383 SNLR---ILGVVKSDGFGYQCVAEAGNAQTSQQLIVPKPAIPSSS-----VLPS 430  
 Qy 233 LP 234  
 Db 431 AP 432  
 RESULT 13  
 KWLS\_HUMAN  
 ID KWLS\_HUMAN STANDARD; PRT; 1914 AA.  
 AC Q15746; O95796; O95797; O95798; O95799; Q14844; Q16794; Q9UBG5;  
 AC Q9UIT9;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Myosin light chain kinase, smooth muscle and non-muscle isoforms  
 DE (EC 2.7.1.117) (MLCK) [Contains: Telokin (Kinase related protein)  
 DE (KCP)].  
 GN MYLK OR MLCK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Umbilical vein endothelial cells;  
 RX MEDLINE=97304466; PubMed=9160829;  
 RA Garcia J.G.N., Lazar V.L., Gilbert-McClain L.I., Gallagher P.J.,  
 RA Verin A.D.;  
 RA "Myosin light chain kinase in endothelium: molecular cloning and  
 RT regulation";  
 RL Am. J. Respir. Cell Mol. Biol. 16:489-494 (1997).  
 RN [2]

RP REVISIONS.  
RA Birukov K.G., Garcia J.G.N.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3A; 3B AND 4).  
RC TISSUE=Umbilical vein;  
RX MEDLINE=99216419; PubMed=10198165;  
RA Lazar V.L., Garcia J.G.N.;  
RT "A single human myosin light chain kinase gene (MLCK; MYLK).";  
RL Genomics 57:256-267(1999).  
RN [4]  
RN REVISIONS (ISOFORM 2).  
RP Birukov K.G., Garcia J.G.N.;  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 923-1914 FROM N.A.  
RC TISSUE=Hippocampus;  
RX MEDLINE=96121365; PubMed=8575746;  
RA Pottier M.-C., Chelot E., Pekarsky Y., Gardiner K., Rossier J.,  
RT Turnell W.G.;  
RT "The human myosin light chain kinase (MLCK) from hippocampus:  
RT cloning, sequencing, expression, and localization to 3qcen-q21.";  
RL Genomics 29:562-570(1995).  
RN [6]  
RP SEQUENCE OF 1614-1914 FROM N.A.  
RC TISSUE=Lung, and Placenta;  
RX MEDLINE=20007838; PubMed=10536370;  
RA Watterson D.M., Schavocky J.P., Guo L., Weiss C., Chlenski A.,  
RA Shinsky V.P., Van Eldik L.J., Haich J.;  
RT "Analysis of the kinase-related protein gene found at human chromosome  
RT 3q21 in a multi-gene cluster: organization, expression, alternative  
RT splicing and polymorphic marker.";  
RL J. Cell. Biochem. 75:481-491(1999).  
RN [7]  
RP SEQUENCE OF 1456-1914 FROM N.A.  
RC TISSUE=Placenta;  
RA Watterson M.D.;  
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR  
CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE  
CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT  
CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS. CENTRAL  
CC DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE  
CC EDENIA FORMATION. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO  
CC CONTROL THE GROWTH INITIATION OF ASTROCYTIC PROCESSES IN CULTURE  
CC AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED  
CC BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT  
CC IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.  
CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin  
CC light-chain] phosphate.  
CC -!- SUBUNIT: ISOFORM TELOKIN BINDS CALMODULIN.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=6;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1; Synonyms=Non-muscle isozyme;  
CC IsoId=Q15746-1; Sequence=Displayed;  
CC Note=the smooth muscle isozyme and telokin are produced by  
CC alternative initiation at Met-923 and Met-1761 of isoform 1;  
CC Name=2;  
CC IsoId=Q15746-2; Sequence=VSP\_004791;  
CC Name=3A;  
CC IsoId=Q15746-3; Sequence=VSP\_004792, VSP\_004794;  
CC Name=3B;  
CC IsoId=Q15746-4; Sequence=VSP\_004791, VSP\_004792, VSP\_004794;  
CC Name=4;  
CC IsoId=Q15746-5; Sequence=VSP\_004792, VSP\_004793;  
CC Name=Del-1790;  
CC IsoId=Q15746-6; Sequence=VSP\_004795;  
CC Event=Alternative initiation;  
CC Comment=3 isoforms, 1/Non-muscle isozyme (shown here),  
CC smooth-muscle isozyme and telokin, are produced by alternative  
CC initiation at Met-1, Met-923 and Met-1761. Telokin has no  
CC catalytic activity;

CC -!- TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOZYMES ARE  
CC EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN  
CC CULTURED ENDOTHELIUM WITH QUALITATIVE EXPRESSION APPEARING TO BE  
CC NEITHER TISSUE- NOR DEVELOPMENT-SPECIFIC. NON-MUSCLE ISOFORM 2 IS  
CC THE DOMINANT SPICE VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN  
CC HAS BEEN FOUND IN A WIDE VARIETY OF ADULT AND FETAL TISSUES.  
CC -!- PFM: MLCK IS PROBABLY DOWN-REGULATED BY PHOSPHORYLATION.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -!- SIMILARITY: Contains 9 immunoglobulin-like C2-type domains.  
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U48959; AAC18423.2; -;  
CC EMBL; AF069601; AAD15921.2; -;  
CC EMBL; AF069602; AAD15922.1; -;  
CC EMBL; AF069603; AAD15923.1; -;  
CC EMBL; AF069604; AAD15924.1; -;  
CC EMBL; X85337; CAA59685.1; -;  
CC EMBL; AF096771; AAD51380.1; -;  
CC EMBL; AF096766; AAD51380.1; JOINED.  
CC EMBL; AF096767; AAD51380.1; JOINED.  
CC EMBL; AF096768; AAD51380.1; JOINED.  
CC EMBL; AF096769; AAD51380.1; JOINED.  
CC EMBL; AF096770; AAD51380.1; JOINED.  
CC EMBL; AF096774; AAD54018.1; -;  
CC EMBL; AF096771; AAD51381.1; -;  
CC EMBL; AF096769; AAD51381.1; JOINED.  
CC EMBL; AF096770; AAD51381.1; JOINED.  
CC EMBL; X90870; CAA62378.1; -;  
CC HSSP; Q83450; IAO6.  
CC Genew; HGNC:7590; MYLK.  
CC MIM; 600922; -;  
CC GO; GO:0004687; F:myosin light chain kinase activity; TAS.  
CC GO; GO:0004648; P:protein amino acid phosphorylation; TAS.  
CC InterPro; IPR003961; FN.III.  
CC InterPro; IPR007110; Ig-like.  
CC InterPro; IPR003006; Ig\_MHC.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR002290; Ser\_thr\_kinase.  
CC Pfam; PF00041; fn3; 1.  
CC Pfam; PF00047; ig; 8.  
CC Pfam; PF00069; pkinase; 1.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC PROSITE; PS00835; IG\_LIKE; 9.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
CC Transferase; Serine/threonine-protein kinase; Calmodulin-binding;  
CC ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;  
CC Alternative initiation; Alternative splicing.  
CC CHAIN 1 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM NON-  
CC MUSCLE ISOZYME.  
CC CHAIN 923 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM  
CC SMOOTH-MUSCLE ISOZYME.  
CC CHAIN 1761 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM  
CC TELOKIN.  
CC INIT\_MET 923 923 FOR ISOFORM SMOOTH-MUSCLE ISOZYME.  
CC INIT\_MET 1761 1761 FOR ISOFORM TELOKIN.  
CC DOMAIN 33 122 IG-LIKE C2-TYPE 1.  
CC DOMAIN 161 249 IG-LIKE C2-TYPE 2.  
CC DOMAIN 414 503 IG-LIKE C2-TYPE 3.  
CC DOMAIN 514 599 IG-LIKE C2-TYPE 4.  
CC DOMAIN 620 711 IG-LIKE C2-TYPE 5.  
CC DOMAIN 721 821 IG-LIKE C2-TYPE 6.  
CC DOMAIN 1098 1186 IG-LIKE C2-TYPE 7.  
CC DOMAIN 1238 1326 IG-LIKE C2-TYPE 8.

```
FT DOMAIN 1343 1413 FIBRONECTIN TYPE-III.
FT DOMAIN 1464 1719 PROTEIN KINASE.
FT DOMAIN 1711 1774 CALMODULIN-BINDING.
FT DOMAIN 1809 1898 IG-LIKE C2-TYPE 9.
FT NP_BIND 1470 1478 ATP (BY SIMILARITY).
FT BINDING 1493 1493 ATP (BY SIMILARITY).
FT ACT_SITE 1585 1585 BY SIMILARITY.
FT DOMAIN 1906 1914 POLY-GLU.
FT DOMAIN 868 895 5 X 28 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 868 895 I-1.
FT REPEAT 896 923 I-2.
FT REPEAT 924 951 I-3.
FT REPEAT 952 979 I-4.
FT REPEAT 980 998 I-5 (INCOMPLETE).
FT DOMAIN 999 1063 6 X 12 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 999 1003 II-1 (INCOMPLETE).
FT REPEAT 1004 1015 II-2.
FT REPEAT 1016 1027 II-3.
FT REPEAT 1028 1039 II-4.
FT REPEAT 1040 1051 II-5.
FT REPEAT 1052 1063 II-6.
FT VARSPLIC 437 506 VSGIPKEVAMFLEGTVPVRQEGSIEVYEDAGSHYLCCLKA
RTSDSGTYSCTASNAQQVSCSWTLQVER -> G (in
isoform 2 and isoform 3B).
FT VARSPLIC 1433 1439 /FTID=VSP_004791.
DEVEVSD -> MKWRQCT (in isoform 3A,
isoform 3B and isoform 4).
FT VARSPLIC 1473 1545 /FTID=VSP_004792.
Missing (in isoform 4).
/FTID=VSP_004793.

Query Match 6.2%; Score 128; DB 1; Length 1914;
Best Local Similarity 20.3%; Pred. No. 0.051;
Matches 86; Conservative 58; Mismatches 149; Indels 130; Gaps 19;

QY 24 PQNATVLKGSQARFNCTVSGW--KLIMWALSDMVLSVRPMPFIINDFTSRYDQGG 81
Db 39 PRNLCKEGATAKEGFRV-RGYPEPQVTH-----RNGQPIITSGRFL---LDCGI 85
QY 82 NPTSEMIHNVPSDSGNIRCSLQNSRLHGSAYLTQV-----MGEL 123
Db 86 RGTFSLVIHAVHEDRGKYTCATNG--SGARQVTVELTVEGSPAKQLGQPVVSKTLGDR 143
QY 124 FI-----PSV-----NLVVAENE-----PCEVCLPSHWTLDPISWELG- 158
Db 144 FSAVETRESIMGECPPKATKLGRRVVVKGQMGREFSKITGRPQ-----PQVTWLKGN 198
QY 159 LLVSHSSYYFVPSPDLQSAVSILALTPQSNGLTLCVATWKSARKSA-----TVN 210
Db 199 VPLQPSARVSVSEKNGMQ-VLEIHGVNQDDVGVTCLVNVGSGKASAEISLQGLDSAN 257
QY 211 LTWIRCPQDTGGGI-----NIPGVLSLPSLGFSLPTWKGVLGLAGTMLLTPCTLTIR 265
Db 258 RSFVRETKATNSDVRKEVTNVISKESLDSL----- 288
QY 266 CCCRRRCGCCNCCRCFCRCRRKGRFRIQFKKSEKTKNETETESNGEN---SGVNS 322
Db 289 -----EAAKSKNCSFPQGGSPPAWANSQPPP-RSKLSEKDSRPTAQTP 336
QY 323 DEQKTTETASLPKSCSSDPQNRSSCGPPHORADQ-----PPRPASHPQASFNLASPE 378
Db 337 VLQKTSSTLQARVQ---PEPRAPGLVLSGSEERKPPRPRPATFTTROPGLGSQD 393
QY 379 KVS 381
Db 394 VVS 396
```

RESULT 14

PTK7\_HUMAN

ID PTK7\_HUMAN STANDARD; PRT; 1070 AA.

AC Q13308; Q13417;

DT 30-MAY-2000 (Rel. 39, Created)

```
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase-like 7 precursor (Colon carcinoma kinase-4)
DE (CKK-4).
GN PTK7 OR CKK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon carcinoma, and Placenta;
RX MEDLINE=96074849; PubMed=7478540;
RA Moesle K., Jallat B., Alves F., Sures I., Plowman G.D., Ullrich A.;
RT "Colon carcinoma kinase-4 defines a new subclass of the receptor
RT tyrosine kinase family.";
RL Oncogene 11:2179-2184(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=97037064; PubMed=8882711;
RA Park S.-K., Lee H.-S., Lee S.-T.;
RT "Characterization of the human full-length PTK7 cDNA encoding a
RT receptor protein tyrosine kinase-like molecule closely related to
RT Chick KLG.";
RL J. Biochem. 119:235-239(1996).
RN [3]
RP SEQUENCE FROM N.A., AND REVISION TO 834.
RA Lee S.-T., Park S.-K., Lee H.-S., Ji A.R., Jung J.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY
CC THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE
CC PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR
CC PROGRESSION MARKER.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, LIVER, PANCREAS,
CC KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID
CC GLAND, OVARY, BRAIN, HEART AND SKELETAL MUSCLE. ALSO EXPRESSED IN
CC ERYTHROLEUKEMIA CELLS. BUT NOT EXPRESSED IN COLON.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
CC -!- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U33635; AAC87565.1; -.
CC EMBL; U40271; AAC50484.2; -.
CC EMBL; AF447176; AAL39062.1; -.
CC EMBL; AF447157; AAL39062.1; JOINED.
CC EMBL; AF447158; AAL39062.1; JOINED.
CC EMBL; AF447162; AAL39062.1; JOINED.
CC EMBL; AF447164; AAL39062.1; JOINED.
CC EMBL; AF447167; AAL39062.1; JOINED.
CC EMBL; AF447170; AAL39062.1; JOINED.
CC EMBL; AF447171; AAL39062.1; JOINED.
CC EMBL; AF447173; AAL39062.1; JOINED.
CC EMBL; AF447174; AAL39062.1; JOINED.
CC EMBL; AF447175; AAL39062.1; JOINED.
CC HSSP; P08631; LADS.
CC Genew; HGNC:9618; PTK7.
CC MIM; 601890; ..
CC GO; GO:0005886; C:plasma membrane; TAS.
CC GO; GO:0005888; C:proteoglycan integral to plasma membrane; TAS.
CC GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. .; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
```

DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002011; RTKinaseII.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00047; ig; 7.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00408; IGC2; 5.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS50835; IG LIKE; 7.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS50010; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; FALSE NEG.  
 KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;  
 KW Immunoglobulin domain; Repeat.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 31 1070 TYROSINE-PROTEIN KINASE-LIKE 7.  
 FT DOMAIN 31 704 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 705 725 POTENTIAL.  
 FT DOMAIN 726 1070 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 31 120 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 128 218 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 225 317 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 303 407 IG-LIKE C2-TYPE 4.  
 FT DOMAIN 412 497 IG-LIKE C2-TYPE 5.  
 FT DOMAIN 503 586 IG-LIKE C2-TYPE 6.  
 FT DOMAIN 578 680 IG-LIKE C2-TYPE 7.  
 FT DOMAIN 796 1066 PROTEIN KINASE; INACTIVE.  
 FT DISULFID 53 101 BY SIMILARITY.  
 FT DISULFID 150 200 BY SIMILARITY.  
 FT DISULFID 246 301 BY SIMILARITY.  
 FT DISULFID 343 391 BY SIMILARITY.  
 FT DISULFID 433 481 BY SIMILARITY.  
 FT DISULFID 524 570 BY SIMILARITY.  
 FT DISULFID 613 664 BY SIMILARITY.  
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 268 283 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 92 92 P -> R (IN REF. 2 AND 3).  
 FT CONFLICT 147 147 K -> T (IN REF. 2 AND 3).  
 FT CONFLICT 207 207 S -> G (IN REF. 2 AND 3).  
 FT CONFLICT 495 496 VL -> RV (IN REF. 2 AND 3).  
 FT CONFLICT 515 515 G -> E (IN REF. 2 AND 3).  
 FT CONFLICT 881 881 E -> G (IN REF. 2 AND 3).  
 FT CONFLICT 969 969 A -> P (IN REF. 2 AND 3).  
 FT CONFLICT 992 992 S -> F (IN REF. 2 AND 3).  
 SQ SEQUENCE 1070 AA; 118260 MW; 47CDF25B8E3698A5 CRC64;

Query Match 6.0%; Score 125; DB 1; Length 1070;  
 Best local similarity 24.8%; Pred. No. 0.043;  
 Matches 59; Conservative 34; Mismatches 109; Indels 36; Gaps 10;

QY 11 AVSGSGNVEIGPQNAVTLKGSQARFNCTVS-QGWKLIMWALSMDVLSVRFMEPIITN 69  
 Db 218 STADESFARVLAQDVVAVRYEAMFHCQSPPPSLQWLFEDFTPTNRSRPHLR 277  
 QY 70 DRFTSRYDQGNFTGEMIHNVPSDSGNIRCSLQNSR-----LHGSAYLTVMQGEFL 124  
 Db 278 ATVFA-----NGSLILTQVRPNAGVIRGICGQGRGPPILEATLHLAEIEDMPLF 328  
 QY 125 IPSVNLVVAENEPCEVTCLPSHWTLPDLSWE-LGL-LVSHSSYYVFPSPDLSQSVIL 182  
 Db 329 ERFVFTAGSEB---RVTCILPPKGLPPSPVWWEHAGVRLPTHGRVY-----QKGHVYL 378  
 QY 183 ALTPQSN-GTLTCTVATKSLKARKSATVNLTV-----IRCPQDTGGGINIPGVLSL 233

Db 379 ANIAESDAGVYCHAA--NLAGQRQDVNITVATVPSTLKKPQDSQLERKPGYLDCL 434  
 RESULT 15  
 PGEM HUMAN  
 ID PGEM HUMAN STANDARD; PRT; 4391 AA.  
 AC P98160; Q16287; Q9H3V5;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Basement membrane-specific heparan sulfate proteoglycan core  
 DE protein precursor (HSPG) (Perlecan) (PLOC).  
 GN HSPG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92112994; PubMed=1730768;  
 RA Kallunki P., Tryggvason K.;  
 RT "Human basement membrane heparan sulfate proteoglycan core protein: a  
 RT 467-kD protein containing multiple domains resembling elements of the  
 RT low density lipoprotein receptor, laminin, neural cell adhesion  
 RT molecules, and epidermal growth factor.";  
 RL J. Cell Biol. 116:559-571(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon, and Skin;  
 RX MEDLINE=92235084; PubMed=1569102;  
 RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;  
 RT "Primary structure of the human heparan sulfate proteoglycan from  
 RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple  
 RT domains homologous to the low density lipoprotein receptor, laminin,  
 RT neural cell adhesion molecules, and epidermal growth factor.";  
 RL J. Biol. Chem. 267:8544-8557(1992).  
 RN [3]  
 RP SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJS1 TYR-1532.  
 RX MEDLINE=20553141; PubMed=11101850;  
 RA Nicole S., Davoine C.-S., Topaloglu H., Cattolico L., Barral D.,  
 RA Beighton P., Ben-Hamida C., Hammouda H., Cruaud C., White P.S.,  
 RA Samson D., Urizbarea J.A., Lehmann-Horn F., Weissbach J.,  
 RA Hentati F., Fontaine B.;  
 RT "Perlecan, the major proteoglycan of basement membranes, is altered in  
 RT patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia)";  
 RL Nat. Genet. 26:480-483(2000).  
 RN [4]  
 RP SEQUENCE OF 1016-1470 FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=91365376; PubMed=1679749;  
 RA Dodge G.R., Kovalszky I., Chu M.L., Hassell J.R., McBride O.W.,  
 RA Yi H.F., Iozzo R.V.;  
 RT "Heparan sulfate proteoglycan of human colon: partial molecular  
 RT cloning, cellular expression, and mapping of the gene (HSPG2) to the  
 RT short arm of human chromosome 1.";  
 RL Genomics 10:673-680(1991).  
 RN [5]  
 RP SEQUENCE OF 890-1396 FROM N.A.  
 RC TISSUE=Fibrosarcoma;  
 RX MEDLINE=92120460; PubMed=1695141;  
 RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,  
 RA Tryggvason K.;  
 RT "Cloning of human heparan sulfate proteoglycan core protein,  
 RT assignment of the gene (HSPG2) to 1p36.1--p35 and identification of  
 RT a BamHI restriction fragment length polymorphism.";  
 RL Genomics 11:389-396(1991).  
 RN [6]  
 RP SEQUENCE OF 1-21 FROM N.A.  
 RX MEDLINE=94052171; PubMed=8234307;  
 RA Cohen I.R., Graessel S., Murdoch A.D., Iozzo R.V.;  
 RT "Structural characterization of the complete human perlecan gene and  
 RT its promoter.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408 (1993).

CC -!- FUNCTION: This protein is an integral component of basement

CC membranes. It is responsible for the fixed negative electrostatic

CC charge and is involved in the charge-selective ultrafiltration

CC properties. It serves as an attachment substrate for cells.

CC -!- SUBUNIT: Purified perlecan has a strong tendency to aggregate in

CC dimers or stellate structures. It interacts with other basement

CC membrane components such as laminin, prolargin and collagen type

CC IV.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.

CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED

CC AND O-LINKED OLIGOSACCHARIDES.

CC -!- DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel

CC syndrome (SUS1) [MIM:255800]; a rare autosomal recessive disorder

CC characterized by permanent myotonia (prolonged failure of muscle

CC relaxation) and skeletal dysplasia, resulting in reduced stature,

CC kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.

CC -!- SIMILARITY: Contains 4 LDL-receptor class A domains.

CC -!- SIMILARITY: Contains 11 laminin EGF-like domains.

CC -!- SIMILARITY: Contains 3 laminin IV domains.

CC -!- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.

CC -!- SIMILARITY: Contains 3 laminin G-like domains.

CC -!- SIMILARITY: Contains 4 EGF-like domains.

CC -!- SIMILARITY: Contains 1 SEA domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X62515; CAA44373.1; -.

DR EMBL; M85289; AAA52700.1; -.

DR EMBL; AL445795; CAC18534.1; -.

DR EMBL; M64283; AAA52699.1; -.

DR EMBL; S76436; AAB21121.2; -.

DR EMBL; L22078; -. NOT\_ANNOTATED\_CDS.

DR PIR; A38096; A38096.

DR HSPG; P00740; IEDM.

DR Sienna-2DPAGE; P98160; -.

DR Genew; HGNC:5273; HSPG2.

DR MIM; 142461; -.

DR MIM; 255800; -.

DR GO; GO:0005206; F.heparin sulfate proteoglycan; TAS.

DR InterPro; IPR000742; EGF\_2.

DR InterPro; IPR001438; EGF\_II.

DR InterPro; IPR006209; EGF\_like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003598; Ig\_c2.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_V.

DR InterPro; IPR000034; laminin\_B.

DR InterPro; IPR002049; laminin\_EGF.

DR InterPro; IPR001791; laminin\_G.

DR InterPro; IPR002172; LDL\_receptor\_A.

DR InterPro; IPR000082; SEA\_domain.

DR Pfam; PF000008; EGF; 4.

DR Pfam; PF00047; Ig; 22.

DR Pfam; PF00052; laminin\_B; 3.

DR Pfam; PF00053; laminin\_EGF; 7.

DR Pfam; PF00054; laminin\_G; 3.

DR Pfam; PF00057; ldl\_recept\_a; 4.

DR Pfam; PF01390; SEA; 1.

DR PRINTS; PR00010; EGFBL00D.

DR ProDom; PD003031; LDLRECEPTOR.

DR SMART; SM00181; EGF; 15.

DR SMART; SM00180; EGF\_Lam; 12.

DR SMART; SM00409; IG; 22.

DR SMART; SM00408; IGC2; 21.

DR SMART; SM00406; IGV; 7.

DR SMART; SM00281; LamB; 3.

DR SMART; SM00282; LamG; 3.

DR SMART; SM00192; LDLa; 4.

DR SMART; SM00200; SEA; 1.

DR PROSITE; PS00022; EGF\_1; 9.

DR PROSITE; PS01186; EGF\_2; 6.

DR PROSITE; PS00835; IG\_Like; 22.

DR PROSITE; PS00025; LAM\_G\_DOMAIN; 3.

DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 11.

DR PROSITE; PS01209; LDLRA\_1; 4.

DR PROSITE; PS00068; LDLRA\_2; 4.

DR PROSITE; PS50024; SEA; 1.

KW Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;

KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;

KW Extracellular matrix; EGF-like domain; Disease mutation.

FT SIGNAL 1 21

FT CHAIN 22 4391

FT DOMAIN 80 194

FT DOMAIN 198 235

FT DOMAIN 284 320

FT DOMAIN 324 360

FT DOMAIN 367 404

FT DOMAIN 405 504

FT DOMAIN 521 530

FT DOMAIN 531 730

FT DOMAIN 731 763

FT DOMAIN 764 813

FT DOMAIN 814 871

FT DOMAIN 879 923

FT DOMAIN 924 933

FT DOMAIN 934 1125

FT DOMAIN 1126 1158

FT DOMAIN 1159 1208

FT DOMAIN 1209 1265

FT DOMAIN 1275 1324

FT DOMAIN 1325 1334

FT DOMAIN 1335 1529

FT DOMAIN 1530 1562

FT DOMAIN 1563 1612

FT DOMAIN 1613 1670

FT DOMAIN 1677 1771

FT DOMAIN 1772 1865

FT DOMAIN 1866 1955

FT DOMAIN 1956 2051

FT DOMAIN 2052 2151

FT DOMAIN 2152 2244

FT DOMAIN 2245 2340

FT DOMAIN 2341 2436

FT DOMAIN 2437 2533

FT DOMAIN 2534 2629

FT DOMAIN 2630 2726

FT DOMAIN 2727 2826

FT DOMAIN 2827 2924

FT DOMAIN 2925 3021

FT DOMAIN 3022 3112

FT DOMAIN 3113 3211

LDL-RECEPTOR CLASS A 1.

LDL-RECEPTOR CLASS A 2.

LDL-RECEPTOR CLASS A 3.

LDL-RECEPTOR CLASS A 4.

IG-LIKE C2-TYPE 1.

LAMININ EGF-LIKE 1 (N-TERMINAL).

LAMININ DOMAIN IV 1 (DOMAIN III A).

LAMININ EGF-LIKE 1 (C-TERMINAL).

LAMININ EGF-LIKE 2.

LAMININ EGF-LIKE 3.

LAMININ EGF-LIKE 4 (INCOMPLETE).

LAMININ EGF-LIKE 5 (N-TERMINAL).

LAMININ EGF-LIKE 5 (C-TERMINAL).

LAMININ EGF-LIKE 6.

LAMININ EGF-LIKE 7.

LAMININ EGF-LIKE 8.

LAMININ EGF-LIKE 9 (N-TERMINAL).

LAMININ EGF-LIKE 9 (C-TERMINAL).

LAMININ EGF-LIKE 10.

LAMININ EGF-LIKE 11.

IG-LIKE C2-TYPE 2.

IG-LIKE C2-TYPE 3.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 5.

IG-LIKE C2-TYPE 6.

IG-LIKE C2-TYPE 7.

IG-LIKE C2-TYPE 8.

IG-LIKE C2-TYPE 9.

IG-LIKE C2-TYPE 10.

IG-LIKE C2-TYPE 11.

IG-LIKE C2-TYPE 12.

IG-LIKE C2-TYPE 13.

IG-LIKE C2-TYPE 14.

IG-LIKE C2-TYPE 15.

IG-LIKE C2-TYPE 16.

IG-LIKE C2-TYPE 17.

Query Match 5.9%; Score 122.5; DB 1; Length 4391;

Best Local Similarity 25.3%; Fred. No. 0.4;

Matches 62; Conservative 28; Mismatches 80; Indels 75; Gaps 12;

QY 6 LTVPEAVGSG--SGNEVIE-GPQNAVTVKGSQARFNCVTSQG-----WKLIHWALSDM 55

Db 3007 VTVPPSEGSSYRLRSRSPVISIDPPSTVQQGDASFKCLIHGGAAPISLEWKRNOLEDN 3066

QY 56 VVLSVRPMEPIITNDRTFSQRYDQGNFTSEMIHNVEPSDSGNIRC-----SLQ 105

Db 3067 VHTS--PNGSIIT-----IVGTRPSNHGTYRCVSNAYGVAQSVV 3104

QY	106	NSRLHGSAYLTQVMGELFI---	PSVNL-VVAENEPEVTCCLPSHWTRLPLDISWEL----	157
Db	3105	NLSVHGPPTVSVLPEGFVWVKVKAVTLECVSAGEPRS----	SARWTRISSTPAKLEQRT	3160
QY	158	-GLLVSHSSYYFVPEPSDLQSAVSIIALTPQSNGLTLCVATWKS	LKARKSATVNLTVIRC	216
Db	3161	YGLMDSH-----	AVLQISSAKPSDAGTYVCLAQNALGTAQKQEVIV-----	3202
QY	217	PQDTG	221	
Db	3203	--DTG	3205	

Search completed: November 13, 2003, 03:31:46  
Job time : 10.3657 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 02:52:59 ; Search time 34.1179 Seconds  
(without alignments)  
2919.536 Million cell updates/sec

Title: US-09-729-264-6  
Perfect score: 2077  
Sequence: 1 MERHLLTVPAGVSGSGNEV.....HPQAFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query %		Match	Length	DB ID	Description
	Score	Match				
1	1694	81.6	315	4	Q9NS15	Q9NS15 homo sapien
2	579.5	27.9	270	11	Q9DBG2	Q9DBG2 mus musculus
3	153	7.4	492	11	Q8C0U8	Q8C0U8 mus musculus
4	142.5	6.9	344	13	Q9DF61	Q9DF61 gallus gall
5	141	6.8	168	11	Q9P9I2	Q9P9I2 mus musculus
6	138	6.6	168	11	Q8CH20	Q8CH20 mus musculus
7	137.5	6.6	315	13	Q9DG15	Q9DG15 gallus gall
8	137.5	6.6	344	13	Q93242	Q93242 gallus gall
9	136.5	6.6	173	11	Q9B4K2	Q9B4K2 mus musculus
10	135	6.5	1496	4	Q92626	Q92626 homo sapien
11	133	6.4	1427	13	Q91562	Q91562 xenopus lae
12	131	6.3	697	11	Q92390	Q92390 mus musculus
13	131	6.3	697	11	Q9DBP0	Q9DBP0 mus musculus
14	129.5	6.2	164	5	Q22048	Q22048 caenorhabdi
15	129	6.2	6632	5	O01761	O01761 caenorhabdi
16	128	6.2	1445	11	Q63155	Q63155 rattus norv

## SUMMARIES

## ALIGNMENTS

## RESULT 1

ID	Q9NS15	PRELIMINARY;	PRT;	315 AA.
AC	Q9NS15;			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	IGSF5 protein (Fragment).			
GN	IGSF5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,			
RA	Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,			
RA	Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,			
RA	Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,			
RA	Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,			
RA	Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,			
RA	Shinrani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,			
RA	Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,			
RA	Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,			
RA	Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,			
RA	Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,			
RA	Lehrach H., Reinhardt K., Yaspo M.L.;			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL163280; CAB90447.1; -			
DR	InterPro; IPR006209; EGF_like.			
DR	InterPro; IPR003159; Ig.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003106; Ig_MHC.			
DR	Pfam; PF00047; Ig; 2.			
DR	SMART; SM00409; IG; 2.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS50835; IG_LIKE; 2.			
FT	NON_TER	1	315	
FT	NON_TER	1	315	

Q17641 caenorhabdi  
O15146 homo sapien  
Q61987 mus musculu  
Q61988 mus musculu  
O57596 gallus gall  
Q8J127 brachydanio  
Q8ay67 brachydanio  
Q66aa2 homo sapien  
Q95qy1 caenorhabdi  
Q8nfaf5 caenorhabdi  
Q8nfaf6 homo sapien  
Q8nfaf8 homo sapien  
Q8nc34 homo sapien  
Q8nb18 homo sapien  
Q8bg33 mus musculu  
Q18238 caenorhabdi  
Q8uiro0 homo sapien  
Q8nda2 homo sapien  
Q8nfaf7 homo sapien  
Q8msn7 drosophila  
Q9jj09 rattus norv  
Q9wt9 drosophila  
Q8n9y9 drosophila  
Q97174 drosophila  
Q8bje2 mus musculu  
Q9vw64 drosophila  
Q8msq0 drosophila  
Q17362 caenorhabdi  
Q8bkg3 mus musculu





```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Basic protein CKIR3.
GN CKIR3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RA Xu X., Bai X., Silvius D., Escalier D., McFarland L., Xu P.-X.;
RT "Ck2 differentially phosphorylate a family of novel spermatid-specific
RT basic nuclear proteins.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF463502; AAO15675.1; -.
SQ SEQUENCE 168 AA; 18957 MW; 8F30D3D27B9BF595 CRC64;

Query Match 6.6%; Score 138; DB 11; Length 168;
Best Local Similarity 32.0%; Pred. No. 9.1e-05;
Matches 39; Conservative 19; Mismatches 50; Indels 14; Gaps 6;

QY 260 CTTITIRCCCRRCGCGNC-CRCRCFCRCRKRGR-----IQFKKSEKENTKETETESG 314
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 55 CSLSRSCCCRCRC-CYRCRCR--CCSRSRFRSRTTLKFFQITEKGESLQRRIRRQ 111
QY 315 NENSGYNSDEQKTETATASLPPKSCGSDPQRNNSCGPHQRADORPPRPASHPOASFNL 374
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 112 LTRSOLELPE--PTWALEPSEITVAFFSHKANVSDP-----EEVPPCLDSDFPFGDL 165
QY 375 AS 376
DQ :|: ||
DB 166 AS 167

RESULT 7
Q9DGI5 PRELIMINARY; PRT; 315 AA.
AC Q9DGI5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CEPU-Se alpha 1 isoform.
GN CEPU-SE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lodge A.P., McNamee C.J., Howard M.R., Reed J.E., Moss D.J.;
RT "Characterisation of CEPU-Se, a secreted isoform of the IGLON family
RT protein CEPU-1.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292936; AAG01879.1; -.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; IGC2; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
KW Immunoglobulin domain.
SQ SEQUENCE 315 AA; 34606 MW; 68C5D27F0DDC6FB2 CRC64;

Query Match 6.6%; Score 137.5; DB 13; Length 315;
Best Local Similarity 26.0%; Pred. No. 0.00023;
Matches 70; Conservative 38; Mismatches 104; Indels 57; Gaps 14;

QY 3 RHLLTPEAVGSGSGNEVI-EGPNATVLKGSQARENCTVSQ-----GW---KLIMWALS 53
DQ :|: || || || || || || || || || || || || || || || || || || || ||
DB 20 RLLFLVPAGVPVRSGDATFPKAMDNVTRVQGESATLRCSVDNRVTRVAVLNRSSILYAGN 79
QY 54 DMVLISVRPMEPIITNDRTSQRDQGNFTSEMIHNHVPDSGNIRCSLQ-----NS 107
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 DKWCLDPRVLLANTKTQYSIQ-----IHDVDVYDEGPPYTCVQTDNHPKTS 126
QY 108 RLHGSAYLTVQVMGELFIPSVNLVAENPEPCEVTCLPSHWTRLPD--ISWELGLLVSHSS 165
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 RVH----LIVQVSPKITEISSDISINEGGNVSITCIA---TGRPDPTIWR---HISPKA 176
QY 166 YFVPEPSDLQSAVSILALTPQSNGLTCTVATWKSLSKARKSATVNLTV-----IRCPQT 220
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 VGFISEDEYLE-----ITGITREQSGEYECSS-NDVAAPVQVRKVTVNPPYISDAKST 231
QY 221 GGINIPGVL-----SSLPSLGFSLPTWGK 245
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 232 GVPVGQKGILMCEASAVPSADFQ---WYK 257

Query Match 6.6%; Score 137.5; DB 13; Length 315;
Best Local Similarity 26.0%; Pred. No. 0.00023;
Matches 70; Conservative 38; Mismatches 104; Indels 57; Gaps 14;

QY 3 RHLLTPEAVGSGSGNEVI-EGPNATVLKGSQARENCTVSQ-----GW---KLIMWALS 53
DQ :|: || || || || || || || || || || || || || || || || || || || ||
DB 20 RLLFLVPAGVPVRSGDATFPKAMDNVTRVQGESATLRCSVDNRVTRVAVLNRSSILYAGN 79
QY 54 DMVLISVRPMEPIITNDRTSQRDQGNFTSEMIHNHVPDSGNIRCSLQ-----NS 107
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 DKWCLDPRVLLANTKTQYSIQ-----IHDVDVYDEGPPYTCVQTDNHPKTS 126
QY 108 RLHGSAYLTVQVMGELFIPSVNLVAENPEPCEVTCLPSHWTRLPD--ISWELGLLVSHSS 165
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 RVH----LIVQVSPKITEISSDISINEGGNVSITCIA---TGRPDPTIWR---HISPKA 176
QY 166 YFVPEPSDLQSAVSILALTPQSNGLTCTVATWKSLSKARKSATVNLTV-----IRCPQT 220
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 VGFISEDEYLE-----ITGITREQSGEYECSS-NDVAAPVQVRKVTVNPPYISDAKST 231
QY 221 GGINIPGVL-----SSLPSLGFSLPTWGK 245
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 232 GVPVGQKGILMCEASAVPSADFQ---WYK 257

```

```

Db 20 RLLFLVPAGVPVRSGDATFPKAMDNVTRVQGESATLRCSVDNRVTRVAVLNRSSILYAGN 79
QY 54 DMVLISVRPMEPIITNDRTSQRDQGNFTSEMIHNHVPDSGNIRCSLQ-----NS 107
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 DKWCLDPRVLLANTKTQYSIQ-----IHDVDVYDEGPPYTCVQTDNHPKTS 126
QY 108 RLHGSAYLTVQVMGELFIPSVNLVAENPEPCEVTCLPSHWTRLPD--ISWELGLLVSHSS 165
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 RVH----LIVQVSPKITEISSDISINEGGNVSITCIA---TGRPDPTIWR---HISPKA 176
QY 166 YFVPEPSDLQSAVSILALTPQSNGLTCTVATWKSLSKARKSATVNLTV-----IRCPQT 220
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 VGFISEDEYLE-----ITGITREQSGEYECSS-NDVAAPVQVRKVTVNPPYISDAKST 231
QY 221 GGINIPGVL-----SSLPSLGFSLPTWGK 245
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 232 GVPVGQKGILMCEASAVPSADFQ---WYK 257

RESULT 8
O93242 PRELIMINARY; PRT; 344 AA.
AC O93242;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CEPU-1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura Y., Shirabe K., Fukushima M., Takeshita M., Tanaka H.;
RT "CEPU-1: an immunoglobulin Superfamily Molecule, Has Cell Adhesion
RT Activity and Shows Dynamic Expression Patterns in Chick Embryonic
RT Spinal Cord.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011810; BAA31514.1; -.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; IGC2; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
KW Immunoglobulin domain.
SQ SEQUENCE 344 AA; 37613 MW; 22CAA8F526A6B57E CRC64;

Query Match 6.6%; Score 137.5; DB 13; Length 344;
Best Local Similarity 26.0%; Pred. No. 0.00026;
Matches 70; Conservative 38; Mismatches 104; Indels 57; Gaps 14;

QY 3 RHLLTPEAVGSGSGNEVI-EGPNATVLKGSQARENCTVSQ-----GW---KLIMWALS 53
DQ :|: || || || || || || || || || || || || || || || || || || || ||
DB 20 RLLFLVPAGVPVRSGDATFPKAMDNVTRVQGESATLRCSVDNRVTRVAVLNRSSILYAGN 79
QY 54 DMVLISVRPMEPIITNDRTSQRDQGNFTSEMIHNHVPDSGNIRCSLQ-----NS 107
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 DKWCLDPRVLLANTKTQYSIQ-----IHDVDVYDEGPPYTCVQTDNHPKTS 126
QY 108 RLHGSAYLTVQVMGELFIPSVNLVAENPEPCEVTCLPSHWTRLPD--ISWELGLLVSHSS 165
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 RVH----LIVQVSPKITEISSDISINEGGNVSITCIA---TGRPDPTIWR---HISPKA 176
QY 166 YFVPEPSDLQSAVSILALTPQSNGLTCTVATWKSLSKARKSATVNLTV-----IRCPQT 220
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 VGFISEDEYLE-----ITGITREQSGEYECSS-NDVAAPVQVRKVTVNPPYISDAKST 231
QY 221 GGINIPGVL-----SSLPSLGFSLPTWGK 245
DQ :|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 232 GVPVGQKGILMCEASAVPSADFQ---WYK 257

```



```
Db 515 -GQYEQAV--NIIGSQKVAHLTVQPRVTPVPASIPSDTIVVEGANVQLP-----CSSQ 566
QY 237 GFSLP--TWKQVGLG--AGTMLLTCTLTJ-----RCCCRRCRG 275
Db 567 GEPEPAITWKGQVQTESGKFHISPEGLTINDVGADAGRYECVARNTIG 618

RESULT 11
Q91562 PRELIMINARY; PRT; 1427 AA.
AC Q91562;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tumor suppressor.
GN XDCCA.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95113183; PubMed=7813784;
RA Pierceall W.E.; Reale M.A.; Candia A.F.; Wright C.V.; Cho K.R.;
RA Fearon E.R.;
RT "Expression of a homologue of the deleted in colorectal cancer (DCC)
RT gene in the nervous system of developing Xenopus embryos.";
RL Dev. Biol. 166:654-665(1994).
DR EMBL; U10986; AAA70168.1; -.
DR HSSP; P40189; IBQU.
DR InterPro; IPR003962; FhIII_subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_VHC.
DR Pfam; PF00041; In3; 6.
DR Pfam; PF00047; Ig; 4.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50835; IG_LIKE; 4.
KW Immunoglobulin domain; Repeat.
SQ SEQUENCE 1427 AA; 156533 MW; 61FEA12C9A674972 CRC64;

Query Match 6.4%; Score 133; DB 13; Length 1427;
Best Local Similarity 25.7%; Pred. No. 0.0044;
Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY 21 IEGPQATVLKGSQARFNCTVSGW--KLIMWALSDMVLSVRPMEPIITNDRPTQRYD 78
Db 243 LQRPNSVVAIEGQDAVLECAVS-GYPTPTIVMVGQD-----EPVPIRTR---KYS 288

QY 79 QGNGFTSEMIHNVPSDSGNIRC--SLONSRLHGSAYLTVQVMGELFIPSVNLVVAENE 136
Db 289 VLGG--SNLLISNVTDGAGYTCVATYKNENTSFSADLTVMVPPQFLNHPANLYAESM 346

QY 137 PCVETCL-----PSHWTRLPDISW-ELGLLVSHSYFVPEPSDLQSAVSILALTPQSNGT 191
Db 347 DIEFECAVSGKPS-----PVRKTKNGEVVIPSDFQIVDGSNLR---ILGLVKSDGY 397

QY 192 LTCVA-----TWKSLKARKSATVNLTVI--RCPQDTGGGINIPGVLS 232
Db 398 YQCIANEAGNIQTYAQLIIPDPAVPSSSILPSAPRDV-----VPVLVSS 442

RESULT 12
Q92290 PRELIMINARY; PRT; 697 AA.
AC Q92290;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type IIB Na/phosphate-cotransporter.
GN SLC34A2 OR NPT2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI, TISSUE=Small intestine;
RX MEDLINE=99045724; PubMed=9826740;
RA Hilfinger H.; Hattenhauer O.; Traebert M.; Forster I.; Murer H.;
RA Biber J.;
RT "Characterization of a murine type II sodium-phosphate cotransporter
RT expressed in mammalian small intestine.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14564-14569(1998).
DR EMBL; AF081499; AAC80007.1; -.
DR MGD; MGI:1342284; SLC34A2.
DR InterPro; IPR003841; Na/Pi_cotranspt.
DR Pfam; PF02690; Na_Pi_cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
SQ SEQUENCE 697 AA; 76286 MW; 839E5CCB0F565265 CRC64;

Query Match 6.3%; Score 131; DB 11; Length 697;
Best Local Similarity 20.3%; Pred. No. 0.0026;
Matches 88; Conservative 53; Mismatches 124; Indels 168; Gaps 22;

QY 5 LITVPAVSGSGNEVIEGPQATVLK-----GQARENCTVS 42
Db 380 LNIIVKLGS-----VLRG-QVATVIKTLNTDPPFFAWLTGYLAILVGAGMTFVQSS 433

QY 43 QGWKLIMWALSDMVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVPSDSGNIRC 102
Db 434 SVFTSAMTPLIGIGVISIERAYPLTLGSMI-----GTTTFAILAALASPGNT--LRS 483

QY 103 SLONSRLHGSAYLTVQVMGELFIPSVNLVVAENPECVTCLPSHWTRLP-----DIS 154
Db 484 LQIALCH-----FFNISGI-----LLWYPIPTFLPRLAKGLNIS 522

QY 155 ----WELGLLVSHSYFVPEPSDLQSAVSILALTPQSNGLTCVATWKS LKARKSATVN 210
Db 523 AKYRW--FAVFLIIEFFVTP-----LTVFLS-----LAGWPLVGVGVPIIL 564

QY 211 LTVI-----RCPQDTGGGINIPGVL---SSLPSLGSFLPTWKGVLGAGTMLLTP 258
Db 565 LLLVLCLRLMLQPRCP-----ILPLKLRDMWFLPLMWHSLKPDNV-----ISLAT 611

QY 259 TCGLTIRCCCCRRRCG-----CNC-CCRCFCCKRKGFRIOFKKSKKTKTKETETE 312
Db 612 TC-FQRRCCCCRCVCCVCMCGCKCKCKCCRD-----QGEBEKEKEQ-----656

QY 313 SGNENSGYNSDEQKTTETATSLPPKSCSESSDPPEQRNSSCGPPHQRAQRP RPASHPOASF 372
Db 657 -----DIPVKASGAFDNAMSKEC-----QDEGKGQV 683

QY 373 NLASPEKVNNTTV 385
Db 684 EVLSMKALSNTTV 696

RESULT 13
Q9DBP0 PRELIMINARY; PRT; 697 AA.
AC Q9DBP0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Solute carrier family 34 (Sodium phosphate), member 2.
GN SLC34A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```



```

RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Waterston R.;
RL      Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR      EMBL; AF003131; AAB54132.2; -.
DR      HSP; P56276; ITLK.
DR      WormPep; C09D1.1; CE30426.
DR      InterPro; IPR000577; FGgy_kin..
DR      InterPro; IPR003961; FN_III.
DR      InterPro; IPR007110; IG-like.
DR      InterPro; IPR003598; IG_C2.
DR      InterPro; IPR003006; IG_MHC.
DR      InterPro; IPR001849; PH_MHC.
DR      InterPro; IPR000219; RhoGEF.
DR      InterPro; IPR001452; SH3.
DR      Pfam; PF00041; fn3; 1.
DR      Pfam; PF00047; IG; 47.
DR      Pfam; PF00169; PH; 1.
DR      Pfam; PF00621; RhoGEF; 1.
DR      Pfam; PF00018; SH3; 1.
DR      SMART; SM00408; IGC2; 23.
DR      SMART; SM00325; RhoGEF; 1.
DR      SMART; SM00326; SH3; 1.
DR      PROSITE; PS00010; DH_2; 1.
DR      PROSITE; PS00933; FGgy_KINASES_1; 1.
DR      PROSITE; PS00835; IG_LIKE; 49.
DR      PROSITE; PS00003; PH_DOMAIN; 1.
DR      PROSITE; PS00002; SH3; 1.
KW      immunoglobulin domain; SH3 domain.
SQ      SEQUENCE 6632 AA; 731665 MW; 2623EDD62960E89 CRC64;

      Query Match
      Best Local Similarity 6.2%; Score 129; DB 5; Length 6632;
      Matches 59; Conservative 41; Mismatches 97; Indels 60; Gaps 10;

Qy      6 LTPPEAVGSG-----SGNE-----VIEQPONATVLKGSQARPNC 40
Db      2038 LIIEQAQDSGKITVEASNEGSSSSAQLTVNPESSTPIVVDGPKSVTIKETETAEPKAT 2097

Qy      41 VSOQW--KLIIWALSDMVILSVRMEPIITNDRETSQRYDQGNFTSEMIIHNVPESDSG 98
Db      2098 IS--GFPAPTQKWTINEKIVESRITTIKEDVYT-----LXISNAKIEQTG 2143

Qy      99 NRCSLQNSRLHGSAYITVQVMBELFTPS-----VNLVWAEINPECVTCI---PSHWTRL 150
Db      2144 TVKVTAQNSAGQDSKQADLKEVPNVKAFKFSQSLTDKVADEGEPLRWNLLEIDGPGSPGT-- 2201

Qy      151 PDISWEH-LGLVSHSSSYFFPEPSDLQSAVSILALTQPSNGTLTCVATWKSLSKARSA 209
Db      2202 -EVSNLNGQLPLTSKDTVQVVDHGDGYHTVIAEAKPEMNGTLTAKAKNAGCETSAGV 2260

Qy      210 NLT-----VIRCPQD 219
Db      2261 TVNGGNKKPEFVQAPQN 2277

```

Search completed: November 13, 2003, 03:33:37  
Job time : 35.4512 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:31:20 ; Search time 12.3761 Seconds  
(without alignments)  
1319.641 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077

Sequence: 1 MERHLTVPEAVGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	7.3	404	US-09-638-649-3	Sequence 3, Appli
2	128.5	6.2	1345	US-08-977-767-3	Sequence 3, Appli
3	128	6.2	1447	US-09-041-886-25	Sequence 25, Appl
4	128	6.2	1447	PCT-US94-05277-2	Sequence 2, Appli
5	127.5	6.1	869	US-08-374-834-16	Sequence 16, Appl
6	127.5	6.1	869	US-08-644-271-29	Sequence 29, Appl
7	127.5	6.1	869	US-09-077-955-33	Sequence 33, Appl
8	127	6.1	332	US-09-062-365-1	Sequence 1, Appli
9	124.5	6.0	340	US-09-651-200-2	Sequence 2, Appli
10	124.5	6.0	441	US-09-651-200-4	Sequence 4, Appli
11	123.5	5.9	534	US-09-651-200-6	Sequence 6, Appli
12	123.5	5.9	534	US-09-651-200-24	Sequence 24, Appl
13	122	5.9	318	US-08-633-148-4	Sequence 4, Appli
14	122	5.9	340	US-08-633-148-2	Sequence 2, Appli
15	119	5.7	352	US-09-996-243-505	Sequence 505, App
16	114	5.5	868	US-08-374-834-1	Sequence 1, Appli
17	114	5.5	868	US-08-644-271-1	Sequence 1, Appli
18	114	5.5	868	US-09-077-955-1	Sequence 1, Appli
19	114	5.5	1395	US-09-540-245A-15	Sequence 15, Appl
20	113.5	5.5	416	US-09-638-649-1	Sequence 1, Appli
21	113	5.4	689	US-09-499-964-1	Sequence 1, Appli
22	111	5.3	365	US-08-979-424-3	Sequence 3, Appli
23	111	5.3	365	US-09-272-436-2	Sequence 2, Appli
24	109.5	5.3	325	US-09-651-200-20	Sequence 20, Appl
25	109.5	5.3	478	PCT-US95-08493-15	Sequence 15, Appl
26	109.5	5.3	860	PCT-US95-08493-19	Sequence 19, Appl
27	109.5	5.3	868	PCT-US95-08493-21	Sequence 21, Appl

28	109.5	5.3	890	1	US-08-445-640-2	Sequence 2, Appli
29	109.5	5.3	890	3	US-08-170-558-2	Sequence 2, Appli
30	109.5	5.3	890	3	US-08-447-314-2	Sequence 2, Appli
31	109.5	5.3	890	3	US-08-445-461-2	Sequence 1, Appli
32	109.5	5.3	911	1	US-08-286-305A-1	Sequence 2, Appli
33	109.5	5.3	911	2	US-08-441-104A-1	Sequence 1, Appli
34	109.5	5.3	911	2	US-08-440-816A-1	Sequence 1, Appli
35	109.5	5.3	911	3	US-09-417-381A-1	Sequence 1, Appli
36	107.5	5.2	362	1	US-08-415-751-6	Sequence 6, Appli
37	106.5	5.1	946	5	PCT-US95-08493-13	Sequence 13, Appl
38	106	5.1	801	1	US-07-906-349A-6	Sequence 6, Appli
39	105	5.1	319	1	US-08-597-495B-22	Sequence 22, Appl
40	105	5.1	319	3	US-09-068-051A-22	Sequence 22, Appl
41	105	5.1	319	4	US-09-336-536-67	Sequence 67, Appl
42	105	5.1	319	4	US-09-254-465A-6	Sequence 6, Appli
43	105	5.1	365	3	US-08-928-383B-2	Sequence 2, Appli
44	105	5.1	690	4	US-08-935-433-2	Sequence 2, Appli
45	105	5.1	690	4	US-09-553-132-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-638-649-3

; Sequence 3, Application US/09638649

; Patent No. 6563015

; GENERAL INFORMATION:

; APPLICANT: Stern, David M.

; APPLICANT: Schmidt, Ann Marie

; APPLICANT: Yan, Shi Du

; TITLE OF INVENTION: TRANSGENIC MICE OVER-EXPRESSING RECEPTOR FOR ADVANCED

; TITLE OF INVENTION: GLYCATION ENDPRODUCT (RAGE) AND MUTANT APP IN BRAIN AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: 0575/62175

; CURRENT APPLICATION NUMBER: US/09/638,649

; CURRENT FILING DATE: 2000-08-14

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patentt Ver. 2.1

; SEQ ID NO 3

; LENGTH: 404

; TYPE: PRT

; ORGANISM: Human

; US-09-638-649-3

Query Match		7.3%	Score 151;	DB 4;	Length 404;
Best Local Similarity		23.3%	Pred. No. 2.3e-05;		
Matches		90;	Conservative	38;	Mismatches 112; Indels 146; Gaps 19;
QY	9	PRAYGSGNEVIRGPONATVLKGSQARFNCTVSGQ---	WKLIIMWALSDMVVLSVRPMEP	65	
Db	124	PEIVDSAS--ELTAGVFN-----KVGTCVSEGSYPAGTLSWHLDG-----	KP	163	
QY	66	IITNDRFTS-----QRYDQGNFT--SEMIHNVPPSDGNIR-----	CSLQNSRLHGSAY	114	
Db	164	LVPKEGVSVKEQRRHPETGLTQLQSELM---VTPARGGDPRTFTSCSFGLPRHRL	220		
QY	115	LTQVMGELFTP---SVNLVVAENP-----	CEVTLCPSHWTRLDPDISWE	156	
Db	221	RTAPIQPRVMEFVPLEEVQLVV---EPGGAVAPGTVTLTCEVPAQS---	PQIHHM	272	
QY	157	LGLVSHSSYFFVPEPSDLQSAVSLALTPOSGNGLTCVATWKSLEKARKSATVNLTVTRC	216		
Db	273	KD-----GVELPLPPSPVLILPEIGPDQGTYSVATSHSHGPOESRAVSISIE-	322		
QY	217	PQDTGGGGINIGVLLSPLSGFSLPTWKVGLGLAGTMLLT-----	PTCTLTIRCCCC	269	
Db	323	PGEEG-----PTAGSVGSGGLGTALALGILGSLGTAALLGVILM	363		
QY	270	RRRCGCGNCCRCFCCKRRKRGFRIOFKKSEKKT--NKETETESGNGNSGYNSDEQKT	327		
Db	364	QRR-----QRRGEERKAPENQEEERAEIN-----	389		

```
QY 328 TETASLPKSCSSDPBORNSSCGPP 353
Db 390 -----QSEEPAGESSTGGP 404

RESULT 2
US-08-977-767-3
; Sequence 3, Application US/08977767
; Patent No. 5972684
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Greenwald, Sara
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/977,767
; APPLICATION NUMBER: 36,749
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0423 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1532042
; US-08-977-767-3

Query Match 6.2%; Score 128.5; DB 2; Length 1345;
Best Local Similarity 33.0%; Pred. No. 0.013;
Matches 37; Conservative 1; Mismatches 41; Indels 33; Gaps 5;

QY 190 GTLTCVATWKLKARKSATVNLTVIRCPQDTGGGI-----NIPGVLSLPSLPSLPTWGWK 245
Db 414 GTCTCTGT-----GC-CGTGGGAGCGTCAGAGCCCGTGGATGTGA 455

QY 246 VGLGLAGTMLLT-PFTLTIRCCRRCCGNCRC-----CFCC 286
Db 456 CGTGAAGAGGTCCTCTATGACCCCTCTCTGCCCCCTCTGAGACTCAGCACC 507

RESULT 3
US-09-041-886-25
; Sequence 25, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bre desen, Dale E.

; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/041,886
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-041-886-25

Query Match 6.2%; Score 128; DB 3; Length 1447;
Best Local Similarity 24.8%; Pred. No. 0.015;
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 14 SGSGNEV-----IEGPQATVLKGSQARFNCTVSQGW--KLIMWALS DM 55
Db 220 SRTGNEAEVRILSDPGLHRQLYFLQPSNVVAIEGDAVLECCVS-GYPPSF T WLRGEE 278

QY 56 VVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHHNVEPSDSGNIRC--SLQNSRLHSGA 113
Db 279 VI-----QLRSKKYSLGG--SNLLISNVTDDSGMYTCVWTKYKNENISASA 323

QY 114 YLTQVMGELFIPSVNLVVAENEPCEVTCLPSHWTRLPDISW-ELGLLVSHSSYFVPEP 172
Db 324 ELTVLPFPFWFLNHPNLYAYESMDIEFECTVS-GKPVPTVMNMKNGDVIPSDYFQIVGG 382

QY 173 SDLSQSAVSIILATPQSNGTTLTCVATWKLKARKSATVNLTVIRCPQDTGGGINIPGVLS 232
Db 383 SNLR---ILGVVKSDEGFYQCVAEAGNAQTSQALIVPKPAIPSSS-----VLPS 430

QY 233 LP 234
Db 431 AP 432

RESULT 4
PCT-US94-05277-2
; Sequence 2, Application PC/TUS9405277
; GENERAL INFORMATION:
; APPLICANT: Bruskin, Arthur
; APPLICANT: Jancos, David E.
; APPLICANT: Johnson, Karen
; APPLICANT: Kinzier, Kenneth W.
; APPLICANT: Vogelstein, Bert
; APPLICANT: Zabrecky, James R.
; TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05277  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagen, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.42709  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9239  
TELEX: 197430 BMB UT  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1447 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-05277-2

Query Match 6.2%; Score 128; DB 5; Length 1447;  
Best Local Similarity 24.8%; Pred. No. 0.015;  
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 14 SGSGNEV-----LEGONATVLKGSQARFNCTVSQGW-KLIMWALSDM 55  
DB 220 SRGNEAEVRLSDPGLHRQLYFLQRFNSVVAIEGKDAVLECCVS-GTPPPSFTWLGRGE 278  
QY 56 VLVSRPMEPIITNDRFTSQRYDQGNFTSEMIHNVPESDSGNIRC--SLQNSRLHGS 113  
DB 279 VI-----QLRSKYSLLGG--SNLLISNVTDSDSGMYTCVVTYKNEISASA 323  
QY 114 YLTQVQMGELFIPSVNLVVAENECVETCLPSHWTRLPDISW-ELGLLVSHSSYFVPEP 172  
DB 324 ELATVLVPPWFLNPSNLAYESMDIBFECTVS-GKPVPTVNMKNGDVIWIPSDYFQIVGG 382  
QY 173 SDLQSAVSILALTPQSNGLTLCVATWKSARKSATVNLTVIRCPQDTGGGINIPGVLS 232  
DB 383 SNLR-----ILGVKSDGFGYQCVAENEAQNSAQLIIVPKAIPSSS-----VLPS 430  
QY 233 LP 234  
DB 431 AP 432

RESULT 5  
US-08-374-834-16  
Sequence 16, Application US/08374834  
Patent No. 5656473  
GENERAL INFORMATION:  
APPLICANT: Valenzuela, et al.  
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/374,834  
FILING DATE: 19-JAN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/095,658  
FILING DATE: 21-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Cobert, Robert J.  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 190A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (914) 345-7400  
TELEFAX: (914) 345-7721  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 869 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-374-834-16

Query Match 6.1%; Score 127.5; DB 1; Length 869;  
Best Local Similarity 20.9%; Pred. No. 0.0084;  
Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

QY 19 EVIEGPONATVLKGSQARFNCTVSQGWK-LIMWALSDMVLSVRPMBPIITNDRFTSQRY 77  
DB 122 KITRPPINVKIIEGLKAVLPCTTMGNPKPSVSWIKGD-----SPLRENSRIAVLE- 171  
QY 78 DQGNFTSEMIHNVPESDSGNIRC--SLQNSRLHGSAY-LTVQVWGMELFIPSVNLVVAENE 136  
DB 172 -----SGSRHNVQEKDAGQYRCVAKNSL--GTAYSKVVKLEVEFARILRAPESHNV 223  
QY 137 P-----CEVTCLPSHWTRLPDISW-ELGLLVSHSSYFVPEPFDIQAQSAVSILALTPQ 187  
DB 224 TFGSFVTLHCTATGIP-----VPTITWIENGNAVSSGSIQESVKDRVIDSLQLFITKP- 277  
QY 188 SNGTLTCVAT-----WKSARKSATVNLTVIRCPQDTGGG-----INIPGVLSLP 234  
DB 278 --GLYTCTATNKGKEFTAKAAATISIAEWSKPKDKNGKGYCAQYRGVCNAVLAKDALV 335  
QY 235 SLGFSL-----PTWGVGLGLAGTMTLTPTCTLTIRCCERRCCGN----- 277  
DB 336 FLNTSYADPBEAQELLVHTAWNEL-----KVVSPVCPRAEAALLCNHIFQECSPGVVP 388  
QY 278 -----CCCRCC-----FCCRRKRGFRIQFKKSEKTKETETESGNENSGYNSDEQKT 327  
DB 389 TPIPCREYCLAVKELFCAKE-----WLVMEKTHRGLYRSEMHLISVPECKLPSMHWDP 444  
QY 328 TETASLPKSCSSDPPQRNNSCGPPHQRADRP-----PRPASHPQASFNIA 375  
DB 445 TACARLP-----HLDYKNENLKTPEP--MTSSKSPSDIPNLPSSSSSSSFSVS 489

RESULT 6  
US-08-644-271-29  
Sequence 29, Application US/08644271  
Patent No. 5814478  
GENERAL INFORMATION:  
APPLICANT: Valenzuela, et al.  
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill Road  
CITY: Tarrytown  
STATE: NY

```
/ COUNTRY: USA
/ ZIP: 10591
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/644,271
/ FILING DATE: 10-MAY-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: USN 60/008,657
/ FILING DATE: 15-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Cobert, Robert J
/ REGISTRATION NUMBER: 36,108
/ REFERENCE/DOCKET NUMBER: REG 195A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 914-345-7400
/ TELEFAX: 914-345-7721
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 869 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-644-271-29

Query Match 6.1%; Score 127.5; DB 2; Length 869;
Best Local Similarity 20.9%; Pred. No. 0.0084;
Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

Qy 19 EVIEGPONATVLKGSQARFNCTVSQGWK-LIMWALSDMVVLSVRPMPPIITNDRFTSQRY 77
Db 122 KITRPPINVKIIIEGLKAVLPCTTMGNPKPSVSIKGD-----SPLRENSRIAVLE- 171
Qy 78 DOGNFTSEMIHNVEPDSGNIRCSLQNSRLHGSAY-LTVQVMGELFIPSVNLVVAENE 136
Db 172 -----SGSLRIHNQVEDAGQYRCVAKNSL--GTAYSKVKVLEVEVFARILRAPESHNV 223
Qy 137 P-----CEVTLCPHSHWTRLPDISW-ELGLLVSHSSYYFVPEPSDLSQSAVSIALTTPQ 187
Db 224 TFGSFVTLHCTATGIP-----VPTITWIENGNAVSSGSIQESVKDRVIDSRLOLFIITKP- 277
Qy 188 SNGTLTCVAT-----WKSLLKARKSATVNLTVIRCPQDTGGG-----INIPGVLSLSP 234
Db 278 --GLYTICIAINKHGEKSTAKAAATISIAEWSKPKQDNKGVCAGYRGEVNAVLAADLV 335
Qy 235 SLGFSL-----PTWKGVLGLAGTMLLTPTCTLTIRCCCRRCRCCGN----- 277
Db 336 FLNTSYADPEEAQELLVHTAWNEL-----KVVSPVCRPAEALLCNHIFQECSPGWVP 388
Qy 278 -----CCCRCC-----FCCRRKRGFRIOFKSEKTKETETESGNSGNSDQKT 327
Db 389 TPIPICREYCLAVKELFCAKE-----WLVMEEKTHRGLYRSEMHLLSVPECSKLPMSHMDP 444
Qy 328 TETASLPKSCSSDPQRNSCGPPHQARDP-----PRPASHPQASFNLA 375
Db 445 TACARLP-----HLDYNKENLKTFFP--MTSSKPSVDIPNLPSSSSSSFSVS 489

RESULT 8
US-09-062-365-1
/ Sequence 1, Application US/09062365
/ Patent No. 6465422
/ GENERAL INFORMATION:
/ APPLICANT: Schmidt, Ann Marie
/ APPLICANT: Stern, David
/ TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A
/ FILE OF INVENTION: SUBJECT
/ FILE REFERENCE: 55424
/ CURRENT APPLICATION NUMBER: US/09/062,365
/ CURRENT FILING DATE: 1998-04-17
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: Patent in Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 332
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-062-365-1

Query Match 6.1%; Score 127; DB 4; Length 332;
Best Local Similarity 25.0%; Pred. No. 0.0024;
Matches 70; Conservative 31; Mismatches 87; Indels 92; Gaps 15;

Qy 9 PEAVGSGSGNEVIEGPNATVLKGSQARFNCTVSQGWK-LIMWALSDMVVLSVRPMP 65
```

Db 102 PEIVDSAS--ELTAGVFN-----KVGTCVSEGSYPAGTILSWHLDG-----KP 141  
Qy 66 IITNDRFTS-----QRVDQGNFT--SEMIHNVPEPSDGNIR-----CSLQNSRLHGSAY 114  
Db 142 LVNFKGVSVKEQTRRPETGLFTLOBELM---VTPARGDPRTFFSCSFPGUPRHRAL 198  
Qy 115 LTVOVMGELFIP-----SVNLVAENEP-----CEVTCILPSSHWRLLPDISWE 156  
Db 199 RTAPIQPRVMEPVPLEEVLVW---EPEGGAVAPGVTTLTCEVPAQPS-----PQIHWM 250  
Qy 157 LGLLVSHSSYFVPEPSDQSAVILALTPOSNGTLTCVATWKSILKARKSATVNLTVIRC 216  
Db 251 KD-----GVPLPLPPSVLILPEIGPODQGTGYSCVATHSHGHGPOESRAVSISIE- 300  
Qy 217 PQDTGGGINIPGVLLSLPSLGLFTWKGVLGLAGTMLL 256  
Db 301 PGBEG-----PTAGSVGGSLGTAL 321

RESULT 9  
US-09-651-200-2  
; Sequence 2, Application US/09651200  
; Patent No. 6429303  
; GENERAL INFORMATION:  
; APPLICANT: Green et al  
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B  
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 15966-562 (CURA-62)  
; CURRENT APPLICATION NUMBER: US/09/651,200  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/152383  
; PRIOR FILING DATE: 1999-09-03  
; PRIOR APPLICATION NUMBER: 60/172909  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/183578  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-651-200-2

Query Match 6.0%; Score 124.5; DB 4; Length 340;  
Best Local Similarity 21.2%; Pred. No. 0.0042;  
Matches 77; Conservative 51; Mismatches 127; Indels 109; Gaps 17;  
Qy 14 SSGSNEVIEGPONATV-LKGSQARFNCTVS--QGWKL-----IMWALSDMVLSVRPMEPI 66  
Db 48 SPTGAVEVQVPEDPVVALVGTDTATLHCSFSPGFSLTQLNLIWQLTDTKQLV----- 100  
Qy 67 ITNDRFTSQRYDQGNF-----TSMIHNHVPSPSGNIRCSLQNSRLHGS 112  
Db 101 ---HSFTEGR-DQGSAYANRTALPDDLAAQNASLRQVRVADEGSTCFV-SIRDFGS 155  
Qy 113 AYLTVQVMGELFIPSVNLV-----VAENPECVTCLPSSHWRLLP--DISWELGL---LVSH 163  
Db 156 AAVSLQVAAPYKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDQGVPLTGN 213  
Qy 164 SSVYFVPEPSDQSAVILALTPOSNGTLTCVATWKSILKARKSATVNLTVIRCP---QDT 220  
Db 214 VTTSQMANEQGLFDVHSLRVVLGANGTYS- - - - -LVRNPFVQQDA 255  
Qy 221 GGGINIPGVLLSLPSLGLFTWKGVLGLAGTMLLAPTCTLTIRCCRRRCGCCNCCC 280  
Db 256 HGSVTITGQPMTPPEAL-----WTVGLSVCLIALLV----- 288  
Qy 281 RCCFCCKRKGFRIFQOKKSEKTKETETESNGNSGYNDSQKTTETASLPKSCS 340  
Db 289 ALAFVVCWRK-----IKQSCBENAGAEQDQ-----EGEGSKTALQPLKHS 331

Qy 341 SDPE 344  
Db 332 KEDD 335  
RESULT 10  
US-09-651-200-4  
; Sequence 4, Application US/09651200  
; Patent No. 6429303  
; GENERAL INFORMATION:  
; APPLICANT: Green et al  
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B  
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 15966-562 (CURA-62)  
; CURRENT APPLICATION NUMBER: US/09/651,200  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/152383  
; PRIOR FILING DATE: 1999-09-03  
; PRIOR APPLICATION NUMBER: 60/172909  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/183578  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 441  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-651-200-4

Query Match 6.0%; Score 124.5; DB 4; Length 441;  
Best Local Similarity 21.2%; Pred. No. 0.0061;  
Matches 77; Conservative 51; Mismatches 127; Indels 109; Gaps 17;  
Qy 14 SSGSNEVIEGPONATV-LKGSQARFNCTVS--QGWKL-----IMWALSDMVLSVRPMEPI 66  
Db 149 SPTGAVEVQVPEDPVVALVGTDTATLHCSFSPGFSLTQLNLIWQLTDTKQLV----- 201  
Qy 67 ITNDRFTSQRYDQGNF-----TSMIHNHVPSPSGNIRCSLQNSRLHGS 112  
Db 202 ---HSFTEGR-DQGSAYANRTALPDDLAAQNASLRQVRVADEGSTCFV-SIRDFGS 256  
Qy 113 AYLTVQVMGELFIPSVNLV-----VAENPECVTCLPSSHWRLLP--DISWELGL---LVSH 163  
Db 257 AAVSLQVAAPYKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDQGVPLTGN 314  
Qy 164 SSVYFVPEPSDQSAVILALTPOSNGTLTCVATWKSILKARKSATVNLTVIRCP---QDT 220  
Db 315 VTTSQMANEQGLFDVHSLRVVLGANGTYS- - - - -LVRNPFVQQDA 356  
Qy 221 GGGINIPGVLLSLPSLGLFTWKGVLGLAGTMLLAPTCTLTIRCCRRRCGCCNCCC 280  
Db 357 HGSVTITGQPMTPPEAL-----WTVGLSVCLIALLV----- 389  
Qy 281 RCCFCCKRKGFRIFQOKKSEKTKETETESNGNSGYNDSQKTTETASLPKSCS 340  
Db 390 ALAFVVCWRK-----IKQSCBENAGAEQDQ-----EGEGSKTALQPLKHS 432  
Qy 341 SDPE 344  
Db 433 KEDD 436

RESULT 11  
US-09-651-200-6  
; Sequence 6, Application US/09651200  
; Patent No. 6429303  
; GENERAL INFORMATION:  
; APPLICANT: Green et al  
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B  
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and

; TITLE OF INVENTION: Polypeptides Encoded Thereby  
 ; FILE REFERENCE: 15966-562 (CURA-62)  
 ; CURRENT APPLICATION NUMBER: US/09/651,200  
 ; PRIOR FILING DATE: 2000-08-30  
 ; PRIOR APPLICATION NUMBER: 60/152383  
 ; PRIOR FILING DATE: 1999-09-03  
 ; PRIOR APPLICATION NUMBER: 60/172909  
 ; PRIOR FILING DATE: 1999-12-21  
 ; PRIOR APPLICATION NUMBER: 60/183578  
 ; PRIOR FILING DATE: 2000-02-18  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6  
 ; LENGTH: 534  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-651-200-6

Query Match 5.9%; Score 123.5; DB 4; Length 534;  
 Best Local Similarity 21.2%; Pred. No. 0.0097;  
 Matches 77; Conservative 50; Mismatches 128; Indels 109; Gaps 17;  
 QY 14 SSGSNEVIEGPONATV-LKGSQARFNCTVS--QGWKL-----IMWALSDMVVLSVRPMEPI 66  
 DB 242 SPTGAVEVQVPEDPVALVGTDLRCFSFPEPGFSLAQLNLINQLTDTKQLV-----294  
 QY 67 ITNDRFTSQRYDQGNF-----TSMIIHNVEPSDGSNIRCSLQNSRLHGS 112  
 DB 295 ---HSFTEGR-DQGSAYANRTALFPDLLAQNASRLQVRVADEGSFTCFV-SIRDFGS 349  
 QY 113 AYLTVQVMGELFIPSVNLV----VAENEPCEVTCPLPSHWTLRP--DISWELGL--LVSH 163  
 DB 350 AAVSLQVAAPYKPSMTLEPNKDLRPGDVTIITC--SSYRGYPAEVFWODGQGVPLTGN 407  
 QY 164 SSYFVPEPSDQSAVSILALTPOSNGTLTCVATWKSARKSATVNLTVIRCP---QDT 220  
 DB 408 VTTSQMANEQGLFDVHSLRVLVGANGTYSC-----LVRNVLQODA 449  
 QY 221 GGGINIPGVLSLSLPSLGSFTWKGVLGLAGTMLLTPTCTLTIRCCRRRCGCCNCCC 280  
 DB 450 HGSVTITGQPMTFPEAL-----WTVGLSVCLIALLV-----482  
 QY 281 RCCFCRRKRGRIQFOKKSEKTKETETESGNSGYNSSDBQKTTETASLPKSCES 340  
 DB 483 ALAFVCRK-----IKQSCBENAGAEDQDQ-----EGGSKTALQPLKHSDS 525  
 QY 341 SDPE 344  
 DB 526 KEDD 529

RESULT 12  
 US-09-651-200-24  
 ; Sequence 24, Application US/09651200  
 ; Patent No. 6429303  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Green et al  
 ; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B  
 ; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and  
 ; TITLE OF INVENTION: Polypeptides Encoded Thereby  
 ; FILE REFERENCE: 15966-562 (CURA-62)  
 ; CURRENT APPLICATION NUMBER: US/09/651,200  
 ; CURRENT FILING DATE: 2000-08-30  
 ; PRIOR APPLICATION NUMBER: 60/152383  
 ; PRIOR FILING DATE: 1999-09-03  
 ; PRIOR APPLICATION NUMBER: 60/172909  
 ; PRIOR FILING DATE: 1999-12-21  
 ; PRIOR APPLICATION NUMBER: 60/183578  
 ; PRIOR FILING DATE: 2000-02-18  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 24  
 ; LENGTH: 534

; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism: Sequence  
 ; OTHER INFORMATION: mz5020.protein from Figure 4.  
 US-09-651-200-24  
 Query Match 5.9%; Score 123.5; DB 4; Length 534;  
 Best Local Similarity 21.2%; Pred. No. 0.0097;  
 Matches 77; Conservative 50; Mismatches 128; Indels 109; Gaps 17;  
 QY 14 SSGSNEVIEGPONATV-LKGSQARFNCTVS--QGWKL-----IMWALSDMVVLSVRPMEPI 66  
 DB 242 SPTGAVEVQVPEDPVALVGTDLRCFSFPEPGFSLAQLNLINQLTDTKQLV-----294  
 QY 67 ITNDRFTSQRYDQGNF-----TSMIIHNVEPSDGSNIRCSLQNSRLHGS 112  
 DB 295 ---HSFTEGR-DQGSAYANRTALFPDLLAQNASRLQVRVADEGSFTCFV-SIRDFGS 349  
 QY 113 AYLTVQVMGELFIPSVNLV----VAENEPCEVTCPLPSHWTLRP--DISWELGL--LVSH 163  
 DB 350 AAVSLQVAAPYKPSMTLEPNKDLRPGDVTIITC--SSYRGYPAEVFWODGQGVPLTGN 407  
 QY 164 SSYFVPEPSDQSAVSILALTPOSNGTLTCVATWKSARKSATVNLTVIRCP---QDT 220  
 DB 408 VTTSQMANEQGLFDVHSLRVLVGANGTYSC-----LVRNVLQODA 449  
 QY 221 GGGINIPGVLSLSLPSLGSFTWKGVLGLAGTMLLTPTCTLTIRCCRRRCGCCNCCC 280  
 DB 450 HGSVTITGQPMTFPEAL-----WTVGLSVCLIALLV-----482  
 QY 281 RCCFCRRKRGRIQFOKKSEKTKETETESGNSGYNSSDBQKTTETASLPKSCES 340  
 DB 483 ALAFVCRK-----IKQSCBENAGAEDQDQ-----EGGSKTALQPLKHSDS 525  
 QY 341 SDPE 344  
 DB 526 KEDD 529  
 RESULT 13  
 US-08-633-148-4  
 ; Sequence 4, Application US/08633148  
 ; Patent No. 5864018  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MORSE, MICHAEL J.  
 ; APPLICANT: NAGASHIMA, MARIKO  
 ; APPLICANT: HOLLANDER, DORIS A.  
 ; TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION  
 ; TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP  
 ; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/633,148  
 ; FILING DATE: 16-APR-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MURPHY ESQ., MATTHEW B.  
 ; REGISTRATION NUMBER: 39,787  
 ; REFERENCE/DOCKET NUMBER: 014618-005600US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 326-2400



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

```
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      5.7%; Score 119; DB 4; Length 352;
Best Local Similarity 20.6%; Pred. No. 0.014;
Matches 92; Conservative 47; Mismatches 127; Indels 180; Gaps 25;

Qy      6 LTVPEAVGSGSNEVIEGPQNAVTLKGSQAFNC--TVS-----QGWKLIMWALS----- 53
Db      22 ITTPE-----EMIE-----KAKGETAYLPCFKTLSPEDQGPLDIEMLISPADNQKV 67

Qy      54 DMVLSVRPEIITNDRFTSQRY-DOGG--NFTSEMI-----IHNVEPSDSGNIRC 102
Db      68 DQVII-----LYSGKIVDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy      103 SLQ-----NSRLHGSAYLTVQVNGELFIPSVNLVVAENBFC-----EVTCLP 144
Db      121 KVKKAPGVANKKIH-----LVVLVKPSGARCVDGSEEIGSDFKIKCEP 164

Qy      145 S-----HWTRLPD-----ISWELGILVSHSSYYFVPEPSDLQSAVSILALTPOSNGT 191
Db      165 KEGSLPQYEWQKLSDSQKMTPTSW-----LAEMTS-----SVISVKNASSEYSGT 209

Qy      192 LTCVATWKSILKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGFSLPTWKGKVL--- 248
Db      210 YSC--TVNRVGSQDCLRLNVV-----PPSNKAGLIAG 241

Qy      249 GLAGTMLLTPTCTLTIRCCCCRRCCGNCRCRKRKGFRIQPKKKEKTN-- 306
Db      242 AIIGTLIALALIGLII-----FCRKKRR-----EEKYEKEVHDI 277

Qy      307 -----KETETESGNSGNSDQKTTETATSLPPKSCSSDPEQNSSCGPPHQRADOR 360
Db      278 REDVPPPKRTSTARSVIGNH-----SSLGMSPSNMEGYSKTQYNQVPSDEFERTPQS 332

Qy      361 PPRPASHPQASFNLASPEKVSNTTW 386
Db      333 P-----TLPPAKFKY--PVKTDGITW 352
```

Search completed: November 13, 2003, 03:35:13  
Job time : 13.3761 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:34:36 ; Search time 26.7591 Seconds  
(without alignments)  
2477.484 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077

Sequence: 1 MERHLLTPEAVGSGGNEV.....HPQASFNLPSPKVSNTTVV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145.5	7.0	405	8	US-08-755-235-4
2	139.5	6.7	344	15	US-10-306-133-3
3	138.5	6.7	2473	15	US-10-184-644-559
4	138.5	6.7	2473	15	US-10-184-634-559
5	138	6.6	633	12	US-10-180-410-26
6	135	6.5	1477	12	US-10-274-583-20
7	135	6.5	1496	12	US-10-021-660-125
8	134	6.5	592	12	US-10-180-410-2
9	134	6.5	594	12	US-10-180-410-12
10	134	6.5	708	12	US-10-180-410-8
11	134	6.5	708	12	US-10-199-672-584
12	134	6.5	708	12	US-10-187-749-584
13	134	6.5	708	12	US-10-194-457-584
14	134	6.5	708	12	US-10-184-642-584
15	134	6.5	708	12	US-10-196-747-584

16	134	6.5	708	12	US-10-173-689-584	Sequence 584, App
17	134	6.5	708	12	US-10-173-690-584	Sequence 584, App
18	134	6.5	708	12	US-10-173-691-584	Sequence 584, App
19	134	6.5	708	12	US-10-173-692-584	Sequence 584, App
20	134	6.5	708	12	US-10-173-694-584	Sequence 584, App
21	134	6.5	708	12	US-10-173-698-584	Sequence 584, App
22	134	6.5	708	12	US-10-173-699-584	Sequence 584, App
23	134	6.5	708	12	US-10-173-707-584	Sequence 584, App
24	134	6.5	708	12	US-10-174-569-584	Sequence 584, App
25	134	6.5	708	12	US-10-174-583-584	Sequence 584, App
26	134	6.5	708	12	US-10-174-587-584	Sequence 584, App
27	134	6.5	708	12	US-10-174-589-584	Sequence 584, App
28	134	6.5	708	12	US-10-174-591-584	Sequence 584, App
29	134	6.5	708	12	US-10-175-736-584	Sequence 584, App
30	134	6.5	708	12	US-10-175-742-584	Sequence 584, App
31	134	6.5	708	12	US-10-175-744-584	Sequence 584, App
32	134	6.5	708	12	US-10-175-745-584	Sequence 584, App
33	134	6.5	708	12	US-10-175-748-584	Sequence 584, App
34	134	6.5	708	12	US-10-175-751-584	Sequence 584, App
35	134	6.5	708	12	US-10-175-754-584	Sequence 584, App
36	134	6.5	708	12	US-10-176-480-584	Sequence 584, App
37	134	6.5	708	12	US-10-176-489-584	Sequence 584, App
38	134	6.5	708	12	US-10-176-754-584	Sequence 584, App
39	134	6.5	708	12	US-10-176-755-584	Sequence 584, App
40	134	6.5	708	12	US-10-176-759-584	Sequence 584, App
41	134	6.5	708	12	US-10-176-920-584	Sequence 584, App
42	134	6.5	708	12	US-10-176-922-584	Sequence 584, App
43	134	6.5	708	12	US-10-176-924-584	Sequence 584, App
44	134	6.5	708	12	US-10-176-984-584	Sequence 584, App
45	134	6.5	708	12	US-10-179-508-584	Sequence 584, App

## ALIGNMENTS

RESULT 1  
US-08-755-235-4  
; Sequence 4, Application US/08755235  
; Publication No. US20030059423A1  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Wu, Jun  
; TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES  
; FILE REFERENCE: 0575/50159  
; CURRENT APPLICATION NUMBER: US/08/755,235  
; CURRENT FILING DATE: 1996-11-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Human  
US-08-755-235-4

Query Match	7.0%;	Score 145.5;	DB 8;	Length 405;
Best Local Similarity	23.5%;	Pred. No. 0.00039;		
Matches	91;	Conservative 38;	Mismatches 111;	Indels 147;
Gaps	20;			
QY	9	PEAVGSGGNEVIEGPQNAVTLKGSQARFNCVTVSOG---WKLIWMALSDMVLSVFPMPER	65	
Db	124	PEIVDSAS--ELTAGVFN-----KVGTCVSEGSYPAGTLSWHLDG-----KP	163	
QY	66	ITINDRFTS-----ORYDQGGNFT--SEMIHNHVPSPDSGNIR-----CSLQNSRLHGSAY	114	
Db	164	LVPNEKGVSVKEQTRHPETGLFTLQSELM---VTPARGDPRPTFSCFSFGLPHRAL	220	
QY	115	LTVMQMGELFTIP-----SVNLVVAENEP-----CEVTCLPSHWTLPDISWE	156	
Db	221	RTAPIQPRVBEFVPLEEVQLVV---EPEGGA VAPGGTVTLTCEVPAQPS-----PQIHMW	272	
QY	157	LGLLVSHSSYFVPPSPDLQSAVSLIATPQSNGLTCTVATWKSILKARKS-ATVNLTVIR	215	



```
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 559
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-184-634-559

Query Match          6.7%; Score 138.5; DB 15; Length 2473;
Best Local Similarity 28.5%; Pred. No. 0.018;
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

Qy 183 ALTPQSGNTLTCTVATKSLKARKATVNLTVIRCPQDTGGGINPGLVLSLPSLGLSPLT 242
Db 2274 AATTGAAGTTTCAATTAATAATTATATGTTCC-----2307

Qy 243 WGVYGLGLAGTMLTPTCTLTIRCCCR-RRCCGNCNCCRCFC 286
Db 2308 -----ATTCTCATCGCCACCCACCCACCCCGCCACCCACC 2344

RESULT 5
US-10-180-410-26
; Sequence 26, Application US/10180410
; Publication No. US20030148382A1
; GENERAL INFORMATION:
; APPLICANT: SUN, CHAO
; APPLICANT: CARULLI, JOHN P.
; APPLICANT: LUKASHIN, ALEXANDER V.
; APPLICANT: KILBURN, DANIEL R.
; TITLE OF INVENTION: PANCREAS NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A097 CIP
; CURRENT APPLICATION NUMBER: US/10/180,410
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: PCT/US01/19904
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,611
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 26
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-180-410-26

Query Match          6.6%; Score 138; DB 12; Length 633;
Best Local Similarity 24.1%; Pred. No. 0.0031;
Matches 54; Conservative 40; Mismatches 100; Indels 30; Gaps 10;

Qy 14 GSGNEVIEGPQATVILKGSQARENCTVSGWKLIMWALSDMVLSVRPMEPIITNDRFT 73
Db 20 AGSPHFLQPEDLVLLGHEARLPCALGAYWGLVQWTKSLGALGGOR-----DLPG 71

Qy 74 SORYDOGGNFTS-----EMIHNVEPSDSGNIRCSLQNSRLHG-SAYLTV-----QVMGE 122
Db 72 WSRVWISGNAANGHDLHIRPVELEDEASVCEQATAGLSRPAQLHLVLPPEAPVLGG 131

Qy 123 LFIPSVNLVAENPECEVTCLPSHWTR-LPDISW-ELGLLVSHSSY--FVPE--PSDLQ 176
Db 132 ---PSVSLVA--GYFANLTCRSRGDARPTPELLFRDGLVDGATFHQTLTKGTGPGSVE 186

Qy 177 SAVSILALTPQSGNTLTCTVATKSLKARKATVNLTVIRCPQDT 220
Db 187 STLTLTFFSHDDGATLVCRARSQALPTGRDTAITSLOYPEVT 230

RESULT 6
```

```
US-10-274-583-20
; Sequence 20, Application US/10274583
; Publication No. US20030138431A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: LRRCAPS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-119C
; CURRENT APPLICATION NUMBER: US/10/274,583
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/357,600
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/361,196
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1477
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-274-583-20

Query Match          6.5%; Score 135; DB 12; Length 1477;
Best Local Similarity 25.0%; Pred. No. 0.018;
Matches 73; Conservative 36; Mismatches 127; Indels 56; Gaps 15;

Qy 13 GSGNEVIEGPQATVILKGSQARENCTVSGWKLIMWALSDMVLSVRPMEPIITNDR 71
Db 337 GSPARPTFVIQPTQNTVELVGESVTECSATGHPPRISWTRGDRTPLPVDRVNITPS-- 394

Qy 72 FTSORYDOGGNFTSEMIHNVEPSDSGNIRCSLQNS--RLHGSAYLTVQWMEGLFIPSVN 129
Db 395 -----GG-----LYIQNVVQDGSGEYACSATNNIDSVHATAFIIVQALPQFTVTPQD 441

Qy 130 LVAENPECEVTCLPSHWTRLPDISW-ELGLLVSHSSYFVPEPSDLQSAVLSILALTPQS 188
Db 442 RVIEGQTVDFQC-EAKGNPPVIAWTKGSQLSVDRHLVLSGTLR--ISGVALHQD- 497

Qy 189 NGTLTCVATWKSLSKARKSATVNLTV-----IRCPQDT-----GGGINIPGLVLSLPSL 236
Db 498 -GQVECCQAV--NIIGSQKVAHLTVQPRVTVFASIPSDTTVEGVANVQLP-----CSSQ 549

Qy 237 GFSLP--TWGKVLGL--AGTMLTPTCTLT-----RCCCRRCRG 275
Db 550 GEPEFAITWNKDGQVVTESGKFKHISPEGFLTINDVGPADAGRYECVAENTIG 601

RESULT 7
US-10-021-660-125
; Sequence 125, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Glytne, Richard
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 1496
; TYPE: PRT
; ORGANISM: Homo sapiens
```



Matches	53;	Conservative	40;	Mismatches	101;	Indels	30;	Gaps	10;
QY	14	SGSGNEVIEGPQNAVLTKGSOARFNCTVTSQCGKLIIMWALSDMWVLVSVRPMEPIITNDRAFT	73						
Db	20	AGPSPHFLQPPEDLLWLLGEEARPLCALGAYGLVQWTKSGLALGGQR-----DLP	71						
QY	74	SQRYDQGNFTS---EMLIHNVPDSQNIKCSLQNSLHG-SAYLTIV-----QWNGE	122						
Db	72	WSRYWISGNAANGQHDHIRVELEDEASYEQATQAGLRSRPQLHVLVPEAPQVLGG	131						
QY	123	LFIPSVNLVVAENPCEVTCLPSHWTR-LPDIWM-ELGLLVSHSYYY--FVPE--PSDLQ	176						
Db	132	--PSVSILVA--GWPNALTCRSGDARTPELLWFRDGVLLDGATFHQTLLKEGTGPSVSE	186						
QY	177	SAVSILALTPOSNGTGLTCVATWKSILKARKSATVNLTVIRCPQDT	220						
Db	187	STLTLPFPHDGDATFVCRARSOALPTGRDTRAITLSLOYPEVT	230						

```

RESULT 11
US-10-199-672-584
; Sequence 584, Application US/10199672
; Publication No. US20030148442A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/199,672
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-672-584

```

```

Db      20  AGPSPHFLQPEDLVLLGEBERRIPCALGATWGLVQWTKSGIALGGQR-----DLPG 71
Qy      74  SQRYDQGNFTS-----EMIHNVPESDGNIRCSLQNSRLHG-SAYLTV-----QVMGE 122
Db      72  WSRWYISGNAANGQHDHLHIRPVELEDEASYESCOATQAGLRSPQAQLHVLVPPEAPQVLGG 131
Qy     123  LFISVNLVVAENPCEVTCLPSWTR-LPDISW-ELGLLVSHSYY--FVPE--PSDLQ 176
Db     132  ---PSVSLVA--GVFANLTCSRGDARPTPELLFRDGVLLDGATFHTLLKEGTFGSVE 186
Qy     177  SAVSILALTQSQNSGTLCVATWKSISKARKSATVNLTIVRCQDT 220
Db     187  STLTLTFSHDDGATFVCARSQALPTGRDRTAILSLQYPEVT 230

RESULT 12
US-10-187-749-584
; Sequence 584, Application US/10187749
; Publication No. US20030153036A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/187,749
; PRIOR FILING DATE: 2002-07-01
; CURRENT APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-749-584

```

QY 14 SGSGNEVIEGPQNATVLKGSQARFNC TVSQGWKLIMWALSDMVVLSVRPMEPIITNDRFT 73

QY 74 SQRVDQGNFTS---EMIIHNVPSDSGNIRCSLQNSRLHG-SAYLTV-----QVMGE 122  
Db 72 WSRWISGNAANGQHDHLPVELEDEASVECOATQAGLSRPAQLHLVLPPEAPQVLGG 131  
QY 123 LFIPSVNLVVAENPECEVTCLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE--PSDLQ 176  
Db 132 ---PSVSLVA--GVPA NLTCRSRGDARPTPELLWFRDGVLLDGATPHQTLLKEGTGSGVE 186  
QY 177 SAVSILALTPQSGNGLTCTVATWKS LKARKSATVNLTVIRCPQDT 220  
Db 187 STLTLPFFSHDDGATFVCRARSQALPTGRDGTATLTLQYPPEVT 230

## RESULT 13

US-10-194-457-584  
; Sequence 584, Application US/10194457  
; Publication No. US20030153037A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C296  
; CURRENT APPLICATION NUMBER: US/10/194,457  
; CURRENT FILING DATE: 2002-07-11  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 584  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-194-457-584  
Query Match 6.5%; Score 134; DB 12; Length 708;  
Best Local Similarity 23.7%; Pred. No. 0.0081;  
Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;  
QY 14 SGSGNEVIEGPQNATVLKGSQARFNCTVSQGWKLIMWLSDMVLSVRPMEPIITNDRFT 73  
Db 20 AGSPHFLOQPEDLVLLGGEARLPCALGAYWGLVQWTKSGLALGGOR-----DLPG 71  
QY 74 SQRVDQGNFTS---EMIIHNVPSDSGNIRCSLQNSRLHG-SAYLTV-----QVMGE 122  
Db 72 WSRWISGNAANGQHDHLPVELEDEASVECOATQAGLSRPAQLHLVLPPEAPQVLGG 131  
QY 177 SAVSILALTPQSGNGLTCTVATWKS LKARKSATVNLTVIRCPQDT 220  
Db 187 STLTLPFFSHDDGATFVCRARSQALPTGRDGTATLTLQYPPEVT 230

## RESULT 15

US-10-196-747-584  
; Sequence 584, Application US/10196747  
; Publication No. US20030162250A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria

QY 123 LFIPSVNLVVAENPECEVTCLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE--PSDLQ 176  
Db 132 ---PSVSLVA--GVPA NLTCRSRGDARPTPELLWFRDGVLLDGATPHQTLLKEGTGSGVE 186  
QY 177 SAVSILALTPQSGNGLTCTVATWKS LKARKSATVNLTVIRCPQDT 220  
Db 187 STLTLPFFSHDDGATFVCRARSQALPTGRDGTATLTLQYPPEVT 230

## RESULT 14

US-10-184-642-584  
; Sequence 584, Application US/10184642  
; Publication No. US20030157635A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C194  
; CURRENT APPLICATION NUMBER: US/10/184,642  
; CURRENT FILING DATE: 2002-06-27  
; PRIOR APPLICATION removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 584  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-184-642-584  
Query Match 6.5%; Score 134; DB 12; Length 708;  
Best Local Similarity 23.7%; Pred. No. 0.0081;  
Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;  
QY 14 SGSGNEVIEGPQNATVLKGSQARFNCTVSQGWKLIMWLSDMVLSVRPMEPIITNDRFT 73  
Db 20 AGSPHFLOQPEDLVLLGGEARLPCALGAYWGLVQWTKSGLALGGOR-----DLPG 71  
QY 74 SQRVDQGNFTS---EMIIHNVPSDSGNIRCSLQNSRLHG-SAYLTV-----QVMGE 122  
Db 72 WSRWISGNAANGQHDHLPVELEDEASVECOATQAGLSRPAQLHLVLPPEAPQVLGG 131  
QY 123 LFIPSVNLVVAENPECEVTCLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE--PSDLQ 176  
Db 132 ---PSVSLVA--GVPA NLTCRSRGDARPTPELLWFRDGVLLDGATPHQTLLKEGTGSGVE 186  
QY 177 SAVSILALTPQSGNGLTCTVATWKS LKARKSATVNLTVIRCPQDT 220  
Db 187 STLTLPFFSHDDGATFVCRARSQALPTGRDGTATLTLQYPPEVT 230

## RESULT 15

US-10-196-747-584  
; Sequence 584, Application US/10196747  
; Publication No. US20030162250A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria

```
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F3430R1C346
; CURRENT APPLICATION NUMBER: US/10/196,747
; CURRENT FILING DATE: 2002-07-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-747-584

Query Match      6.5%; Score 134; DB 12; Length 708;
Best Local Similarity 23.7%; Pred. No. 0.0081;
Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;

QY      14  SSGSNEVIEGPQNAVTLKGSQARFNCTVSGQWKLIIMWALSDMVVLSVRPMEPIITNDRFT 73
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      20  AGPSPHLQOPEDLVLLGGEARLPCALGAYGVLQWTKSGLALGGQ-----DLPG 71

QY      74  SQRYDQGNFTS---EMIIHNVPSDSGNIRCSLQNSRLHG-SAYLTV-----QVNGE 122
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      72  WSRYWISGNAANGQHDHIRPVELEDEASVECCATQAGLRSRPAQLHVLVPPEAPQVLGG 131

QY      123  LFIPSVNLVVAENEPCEVTCLPSHWTR-LPDISW-ELGLIVSHSSYY--FVPE--PSDLQ 176
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      132  ---PSVSLVA--GVPANLTCRSQDARPTPELLWFRDGVLLDGATFHQTLLKEGTFGSVE 186

QY      177  SAVSILALTPQSNGLTCTVATWKS LKARKSATVNLTVIRCPQDT 220
      |:::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      187  STLTLPFSDHGDGATFVCRARSQALPTGRDTAITLSLQYPPEVT 230
```

Search completed: November 13, 2003, 03:46:16  
Job time : 27.7591 secs

Matches 702; Conservative 0; Mismatches 354; Indels 69; Gaps 4;	
QY 115	AGTTCTGGTCTGTTAAAGTCTATAGAGCCGCCCAAGATGCAACAGTCTCTGAAGG 174
Db 109	AGTTCCGATCCAGTATCAGATCATAGAGTCTCTCAGATGTAAAGG 168
QY 175	CTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGCTGGAAGTCTCATATGCGGCTCT 234
Db 169	CTCAGAGGCTCACTTCAACTGCAACCGTCTCAAGCTGGAAGCTTCTCATGTGACTCT 228
QY 235	CAGTGACATGCTGCTCAAGCTGAGGCCATGAGGCCATCATCAACATGACCGCTT 294
Db 229	TAACCAAAATGGTGTCTGAGTCTACCAACCAAGGACCCATCATCAACAAACCGCTT 288
QY 295	CACCTCTCAGAGGTAGCACAGGCGGGAACTTCACTCGAGATGATCATCCACAATGT 354
Db 289	CACCTATGCCAGTTACAACAGCACTGACAGCTTCACTTCGGAGTTGATCATCATGATGT 348
QY 355	GGAGCCAGTATTCGGGGAACATCAGATGACGCTCCAGAAACAGTCGCTGATGGATC 414
Db 349	GCAGCCAGTACTCGGGATCCGTGCAATGACGCTGCAGAACAGCCATGGGTTGGATC 408
QY 415	TGCTTACCTTACCGTCCAGTATGGAGAGCTGTTCACTCCAGTGTAAATCTGTAGT 474
Db 409	TGCTTCTCTCAGTCAAGTCAATGGGACCCCTGAACATTCCTAGCAACACCTTATAGT 468
QY 475	CGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCCCTCACACTGGACCGGCTCCCGGA 534
Db 469	CACCTGAGGTGAACCTGTAAATGAGCTTCTATGCGCTGGGCTGAGACCTCACTCCCGGA 528
QY 535	TATTTCTGGGAGCTGGTCTCTGGTACGCAATTCAGATATATTTTGTTCGGAGCC 594
Db 529	TATTTCTGGGAGCTGGTCTCTGGTAAAGCTTGCAGGCTGCAAGTCTTCACTTTCGGAGCC 588
QY 595	CAGCGACCTTCAAGTGCAGTACGATCCTGGCTCTGACCCCAAGCAAGCAATGGGACTTT 654
Db 589	GGGCACTTTATGAGGTCTTGAAGTCTCTGGACCTCACACCACTGGGCAACGGGACCTT 648
QY 655	GACTTGGTGGTCACTTGGAGAGCTTGAAGGCGGCAAGTCTGCAACTGTAATCTCAC 714
Db 649	GACTTGGTGGCAGAGTGAAGGACTTGCAGGCGCAGCAAGTCTTAACTGTCAACCTGAC 708
QY 715	TGTGATTCGTTGCTCCCAAGCACTGAGGTGTTAATATATTCAGGTGATTAATCAAG 774
Db 709	TGTGGTTCAGCTCCACTCACAGTATTTGGAGAGGA----- 744
QY 775	TTTACCGAGTTTAAAGTTTTCATGCTCTACTTGGGCAAAAGTTGGACTTGGACTAGCAGG 834
Db 745	-----AGGCCAGCACTGCGGACCTGGGCCATCATCTGCTGGCAGTGGCTT 792
QY 835	CACCATGCTTCTGAGCGGACGTGTACTCTTACAAATACGCTGCTGCTGCGCGCGTCG 894
Db 793	TTCTTGTCTTGTATCTGATCATCTTGTGTTGATTAATATTTCTGTTGCTG----- 843
QY 895	TTGTTGGTGGTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
Db 844	-----TTGTGCTTCCAGAGAGAAAGAGAAATCTACTTATCAAAATGAATAGGAAT 899
QY 955	TCGTATTCAATTTCAAAAGAAATCTGAAAAAGAGAGACAAACAAAGAAACTGAGACAGA 1014
Db 900	CTGCAACATGAGGACAA-----ACAAAGCAGATCCGGAGACAAAGTTAAA 945
QY 1015	AAGTGGAAATGAATACTCCGGCTACAAATTCAGATGAACAAAGACCAAGACCGGCTTC 1074
Db 946	AAGTGGAAAGAAATACCGGTACAGTTCCGATGAGGCAAAAGGCTGCACAGACTGATC 1005
QY 1075	TCTCCCTCCCAAAATCTCTGTAATCCAGTGTACCTGGAACAAAGAAACAGTGTGCGCC 1134
Db 1006	TCTCCCTTCTTAATCTCTGTAAGTACAGCTTCCAGAAAAACGACAGTA-----GCCT 1059
QY 1135	TCCTCACCGGGCTGATCAACGTTCCACCGGCGCAGCAAGTCAATCCAGGCTTCTTT 1194
Db 1060	TCCTTATCAGGAATCAATAAACATCAGCCCGTCCAGCAACTCATCCAGGGTTTCCTT 1119

QY 1195	TAATCTGGCCAGTCTCTGAGAGGTCAAGTAATCAACTGTAGTATA 1239
Db 1120	TGACATCGCCAGTCTCTCAGAGGTGAGAAATGTGACTTTAGTGTA 1164
RESULT 7	
AAI36582	
ID	AAI36582 standard; DNA; 401 BP.
XX	XX
AC	AAI36582;
XX	XX
DT	17-OCT-2001 (first entry)
XX	XX
DE	Probe #5268 used to measure gene expression in human placenta sample.
XX	XX
KW	Probe; microarray; human; placenta; antenatal diagnosis;
KW	genetic disorder; ss.
OS	Homo sapiens.
XX	XX
PN	WO200157272-A2.
XX	XX
PD	09-AUG-2001.
XX	XX
PF	30-JAN-2001; 2001WO-US00663.
XX	XX
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	XX
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	XX
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	XX
DR	WPI; 2001-488897/53.
XX	XX
PT	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -
XX	XX
PS	Claim 25; SEQ ID No 5268; 654pp; English.
XX	XX
CC	The present invention relates to single exon nucleic acid probes (SENPs).
CC	The present sequence is one such probe. The probes are useful for
CC	producing a microarray for predicting, measuring and displaying gene
CC	expression in samples derived from human placenta. The probes are useful
CC	for antenatal diagnosis of human genetic disorders.
XX	XX
SQ	Sequence 401 BP; 105 A; 98 C; 111 G; 87 T; 0 other;
Query Match 15.2%; Score 188.8; DB 22; Length 401;	
Best Local Similarity 99.8%; Pred. No. 4.2e-42;	
Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY 115	AGGTTCTGGTCTGTTAAAGTCAAGAGGCCGCCAGAAATGCAACAGTCTCTGAAGGG 174
Db 210	AGGTTCTGGTCTGTTAAAGTCAAGAGGCCGCCAGAAATGCAAGAGTCTCTGAAGGG 269
QY 175	CTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGCTGGAAGCTCATATGTTGGGCTCT 234
Db 270	CTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGCTGGAAGCTCATATGTTGGGCTCT 329
QY 235	CAGTGACATGCTGCTCAAGCTCAGGCCCATGAGGCCCATCATCAATGACCGCTT 294
Db 330	CAGTGACATGCTGCTCAAGCTCAGGCCCATGAGGCCCATCATCAATGACCGCTT 389
QY 295	CACCTCTCAGAG 306
Db 390	CACCTCTCAGAG 401

are also useful as immunosuppressive agents for bone marrow and organ transplantation or to prolong graft survival. B7-L molecules are also useful for diagnosis and treatment of diseases involving abnormal cell proliferation, including arteriosclerosis and vascular restenosis. Antagonists of B7-L polypeptides are useful for alleviation of toxic shock syndrome or allosensitisation due to blood transfusions, and for treatment of allergy, asthma and hypersensitivity reactions. Nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and lymphoproliferative disorders such as multiple myeloma. The present sequence represents the coding sequence of mouse B7-L\_m3.

Sequence 754 BP; 220 A; 191 C; 175 G; 168 T; 0 other;

Query Match 15.0%; Score 185.6; DB 24; Length 754;  
Best Local Similarity 73.8%; Pred. No. 4.5e-41;  
Matches 236; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 115 AGTTCTGGGCTGGTAATGAAGTCATAGAGGCCGCCAGATGCAACAGCTCTCTGAAGGG 174  
DB 109 AGCTTCGGATCCAGTTATCAGATCATAGAGGTCCTCAGATGTAACAGTCTCTAAAGGA 168  
QY 175 CTCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGAGCTCATCATGTGGGCTCT 234  
DB 169 CTCAGAGGCTCACTTCAACTGCACCGCTGACTCAGCGGTGGAAGCTTCTCATGTGGACTCT 228  
QY 235 CAGTGACATGCTGCTGAAGCGTCAGGCCCATGAGCCCATCATCAACCAATGACCGCTT 294  
DB 229 TAACCAATGCTGCTGCTGAGTCTCACCACCAAGGACCCATCATCAACCAACCGCTT 288  
QY 295 CACTCTCAGAGGTACGACCCAGGGCGGGAATCTTCACTCGGAGATGATATCCACAATGT 354  
DB 289 CACTATGCCAGTTACAAACAGCAGCTGACAGCTTCACTCGGAGTTGATCATCCATGATGT 348  
QY 355 GGAGCCAGTGATTCCGGGGAACATCAGATGACGCTCCAGAACAGTGCCTGATGGATC 414  
DB 349 GCAGCCAGTGACTCGGGATCCGTCGAATGACGCTGCAGAACAGCCATGGTTGGATC 408  
QY 415 TGCTTACCTTACCGTCCAG 434  
DB 409 TGCTTCTCTCAGTGCAG 428

RESULT 9  
ABK13033  
ID ABK13033 standard; cDNA; 895 BP.  
XX  
AC ABK13033;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE DNA encoding mouse B7-L-like protein, B7-L\_m2.  
XX  
KW Mouse; B7-L-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductiv disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 53..865  
FT /\*tag= a  
FT /product= "B7-L-like protein, B7-L\_m2"  
XX  
PN W0200200710-A2.

are also useful as immunosuppressive agents for bone marrow and organ transplantation or to prolong graft survival. B7-L molecules are also useful for diagnosis and treatment of diseases involving abnormal cell proliferation, including arteriosclerosis and vascular restenosis. Antagonists of B7-L polypeptides are useful for alleviation of toxic shock syndrome or allosensitisation due to blood transfusions, and for treatment of allergy, asthma and hypersensitivity reactions. Nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and lymphoproliferative disorders such as multiple myeloma. The present sequence represents the coding sequence of mouse B7-L\_m3.

Sequence 754 BP; 220 A; 191 C; 175 G; 168 T; 0 other;

Query Match 15.0%; Score 185.6; DB 24; Length 754;  
Best Local Similarity 73.8%; Pred. No. 4.5e-41;  
Matches 236; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 115 AGTTCTGGGCTGGTAATGAAGTCATAGAGGCCGCCAGATGCAACAGCTCTCTGAAGGG 174  
DB 109 AGCTTCGGATCCAGTTATCAGATCATAGAGGTCCTCAGATGTAACAGTCTCTAAAGGA 168  
QY 175 CTCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGAGCTCATCATGTGGGCTCT 234  
DB 169 CTCAGAGGCTCACTTCAACTGCACCGCTGACTCAGCGGTGGAAGCTTCTCATGTGGACTCT 228  
QY 235 CAGTGACATGCTGCTGAAGCGTCAGGCCCATGAGCCCATCATCAACCAATGACCGCTT 294  
DB 229 TAACCAATGCTGCTGCTGAGTCTCACCACCAAGGACCCATCATCAACCAACCGCTT 288  
QY 295 CACTCTCAGAGGTACGACCCAGGGCGGGAATCTTCACTCGGAGATGATATCCACAATGT 354  
DB 289 CACTATGCCAGTTACAAACAGCAGCTGACAGCTTCACTCGGAGTTGATCATCCATGATGT 348  
QY 355 GGAGCCAGTGATTCCGGGGAACATCAGATGACGCTCCAGAACAGTGCCTGATGGATC 414  
DB 349 GCAGCCAGTGACTCGGGATCCGTCGAATGACGCTGCAGAACAGCCATGGTTGGATC 408  
QY 415 TGCTTACCTTACCGTCCAG 434  
DB 409 TGCTTCTCTCAGTGCAG 428

RESULT 8  
ABK13034  
ID ABK13034 standard; cDNA; 754 BP.  
XX  
AC ABK13034;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE DNA encoding mouse B7-L-like protein, B7-L\_m3.  
XX  
KW Mouse; B7-L-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductiv disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 53..724  
FT /\*tag= a  
FT /product= "B7-L-like protein, B7-L\_m3"  
XX  
PN W0200200710-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 28-JUN-2001; 2001WO-US20719.  
XX  
PR 28-JUN-2000; 2000US-214512P.  
PR 28-NOV-2000; 2000US-0729264.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
XX  
DR WPI; 2002-130881/17.  
DR P-PSDB; AAU75546.  
XX  
XX New B7-L-like polypeptides, polynucleotides and their modulators, useful  
XX for diagnosing, preventing and treating reproductive, immune and  
XX proliferative disorders, e.g. cancer and arteriosclerosis -  
XX  
XX Claim 1; Fig 7; 135pp; English.  
XX  
XX The invention relates to an isolated B7-L-like (B7-L) polypeptide (I).  
XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
XX useful for treating B7-L-like polypeptide-related disease, disorders or  
XX conditions including reproductive disorders (e.g. infertility,  
XX miscarriage, preterm labour and delivery and endometriosis) and  
XX proliferative disorders. Antibodies, soluble proteins comprising  
XX extracellular domains and other regulators of B7-L polypeptides are  
XX useful for enhancing the immune response to tumours. (I) plays a role in  
XX growth and maintenance of cancer cells based on the observation of  
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
XX polypeptide. Hence modulators of (I) are useful for the treatment of  
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
XX in allograft transplantation, graft versus host disease, T-cell  
XX dependent B-cell mediated diseases and autoimmune diseases. B7-L  
XX molecules are useful for alleviating the symptoms associated with  
XX diseases involving chronic immune cell dysfunction or to treat  
XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
XX purpura and psoriasis, chronic inflammatory disease such as  
XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They

```

PD 03-JAN-2002.
XX
XX
XX 28-JUN-2001; 2001WO-US20719.
XX
XX 28-JUN-2000; 2000US-214512P.
PR 28-NOV-2000; 2000US-0729264.
XX
XX (AMGE-) AMGEN INC.
XX
XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
XX WPI; 2002-130881/17.
DR P-PSDB; AAU75545.
XX
XX New B7-like polypeptides, polynucleotides and their modulators, useful
PT for diagnosing, preventing and treating reproductive, immune and
PT proliferative disorders, e.g. cancer and arteriosclerosis -
XX
XX Claim 1; Fig 6; 135pp; English.
XX
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).
CC The polypeptide, polynucleotide encoding it and antibody against (I) are
CC useful for treating B7-like polypeptide-related disease, disorders or
CC conditions including reproductive disorders (e.g. infertility,
CC miscarriage, preterm labour and delivery and endometriosis) and
CC proliferative disorders. Antibodies, soluble proteins comprising
CC extracellular domains and other regulators of B7-L polypeptides are
CC useful for enhancing the immune response to tumours. (I) plays a role in
CC growth and maintenance of cancer cells based on the observation of
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
CC polypeptide. Hence modulators of (I) are useful for the treatment of
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
CC in allograft transplantation, graft versus host disease, T-cell
CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
CC molecules are useful for alleviating the symptoms associated with
CC diseases involving chronic immune cell dysfunction or to treat
CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
CC purpura and psoriasis, chronic inflammatory disease such as
CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
CC are also useful as immunosuppressive agents for bone marrow and organ
CC transplantation or to prolong graft survival. B7-L molecules are also
CC useful for diagnosis and treatment of diseases involving abnormal cell
CC proliferation, including arteriosclerosis and vascular restenosis.
CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
CC shock syndrome or allosensitisation due to blood transfusions, and for
CC treatment of allergy, asthma and hypersensitivity reactions,
CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
CC lymphoproliferative disorders such as multiple myeloma. The present
CC sequence represents the coding sequence of mouse B7-Lm2.
XX
XX Sequence 895 BP; 250 A; 223 C; 212 G; 210 T; 0 other;
SQ
Query Match 15.0%; Score 185.6; DB 24; Length 895;
Best Local Similarity 73.8%; Pred. No. 4.9e-41;
Matches 236; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 115 AGGTTCTGGTCTGGTAAAGTCAATAGAGGCCCCAGAAATGCAACAGTCTGAAGG 174
D 109 AGCTTCGGATCCAGTTATCAGATCATAGAGAGGTCCTCAGAAATTAACAGTCTTAAGGA 168
QY 175 CTCCAGGCTCGTTCACCTGACCGTCTCCAGGCTGGAGCTCATCATGTGGGTCT 234
D 169 CTCAGAGCTCACTTCAACTGACCGTACATCAGGCTGAAGTCTTCATGTGGACTCT 228
QY 235 CAGTGACATGGTGTGCTAAGCGTGCAGGCCCATGAGCCCATCATCAACATGACCGCTT 294

```

229 TAACCAAAATGGTGTCTGTGAGTCTCACACCCAAAGGACCCATCATCAACCAACCGCTT 288

295 CACCTCTCAGAGGTACGACACGCGGGGAACTTTCACCTCGGAGATGATCATCCACAATGT 354

289 CACCTATGCCAGTTACAAACAGCACTGACAGCTTCATCTCGAGTTGATCATCCATGATGT 348

355 GGAGCCCAAGTATTCGGGGAACATCAGATGACAGCTCCAGAAACAGTCCGCTGCATGATGC 414

349 GCAGCCCAAGTACTCGGGATCCGTGCAATGACAGCTGCAAGACAGCCATGGGTTTGATC 408

415 TGCTTACCTTACCGTCCAG 434

409 TGCCTTCTCTCAGTCAAG 428

RESULT 10

ABX55016

ID ABX55016 standard; cDNA; 398 BP.

XX AC ABX55016;

XX XX

DT 26-FEB-2003 (first entry)

XX XX

DE Bovine EST associated with lactation/muscle/fat deposition #4945.

XX XX

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

KW muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX XX

OS Bos Taurus.

XX XX

PN US2002137160-A1.

XX XX

PD 26-SEP-2002.

XX XX

PF 26-OCT-2001; 2001US-0983965.

XX XX

PR 17-DEC-1998; 98US-113678P.

PR 15-DEC-1999; 99US-0465231.

XX XX

PA (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

XX XX

PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-102386/09.

XX XX

PT Purified nucleic acid molecules, useful for genome mapping, gene

PT identification and analysis, cattle breeding or preparation of

PT constructs for cattle gene expression and genetically improved cattle -

XX XX

PS Claim 2; SEQ ID No 4945; 38pp; English.

XX XX

CC The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMFD), derived

CC from cattle, and the LMFD nucleic acid can specifically hybridise to a

CC second nucleic acid molecule comprising any of 5912 nucleotide

CC sequences, appearing as ABX50072-ABX55983, or complements of them.

CC Also included are; (1) a transformed cell having a nucleic acid

CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-

CC translated sequence that functions in the cell to cause termination of

CC transcription and addition of polyadenylated ribonucleotides to a 3' end

CC of the mRNA molecule; and (2) determining a level or pattern of a

CC molecule in a bovine cell or tissue comprising: (a) incubating a marker

CC nucleic acid (comprising any of the 5912 nucleic acid sequences or its

CC complement or fragment) with a complementary nucleic acid molecule

CC obtained from the bovine cell or tissue, where hybridisation between the

CC marker nucleic acid and the complementary nucleic acid permits the

CC detection of the molecule; and (b) detecting the level or pattern of the

CC complementary nucleic acid, where the detection of the complementary

CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 5912 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?docID=20020137160.  
XX  
XX  
SQ Sequence 398 BP; 88 A; 102 C; 137 G; 71 T; 0 other;  
Query Match 11.3%; Score 140.4; DB 25; Length 398;  
Best Local Similarity 78.5%; Pred. No. 1.2e-28;  
Matches 168; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
Qy 113 GTAGTTCTGGTCTGGTAATGAATCATAGAGGCCCCAGAAATGCAACAGTCTCTGAAG 172  
Db 185 GCAGCTGTGGATCCAGCAGTGAATCATAGAGGGTCCCAAGATGTCAAGCCCTGAAG 244  
Qy 173 GGCTCCAGGCTCGTTCAACTGACCTCTCCAGGGCTGGAAGCTCATCATGTGGCT 232  
Db 245 GGCTCGAGGGCTCGTTCAACTGACCATCTCGCAGGGCTGGAAGCTCATGTGGCT 304  
Qy 233 CTCACTGACATGTGGTCTGAAGCTCAGGCCATGGAGCCCATCATCAACCAATGACCGC 292  
Db 305 CTGAGAGGCACAGTGTGTGCTGAGCATGACACTATGAGACATCATCAACAGTGACCGC 364  
Qy 293 TTCACCTCTCAGAGTACGACAGGGCGGGAAT 326  
Db 365 TTCACTTCGCAAGCTACCAAGAGGGCGGGAAT 398  
RESULT 11  
ID ABK10239/c  
XX ABK10239 standard; DNA; 277 BP.  
XX AC ABK10239;  
XX  
XX  
DT 20-MAY-2002 (first entry)  
XX  
XX Trinucleotide repeat sequence #2.  
XX Trinucleotide repeat; fragile X syndrome; ds; DRPLA;  
KW spinocerebellar ataxia type III; Marfan syndrome;  
KW hereditary hypertrophic cardiomyopathy; neuropsychiatric;  
KW dentatorubral and pallidolysian atrophy;  
KW spinocerebellar ataxia; X-linked spinobular atrophy.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT repeat\_region 152..194  
FT /\*tag= a  
FT /rpt\_type= TANDEM  
FT repeat\_unit 152..154  
FT /\*tag= b  
FT /note= "CAG type repeat"  
FT repeat\_region 198..222  
FT /\*tag= a  
FT /rpt\_type= TANDEM  
FT repeat\_unit 198..200  
FT /\*tag= b  
FT /note= "CAG type repeat"  
XX  
XX Homo sapiens.  
XX KR200003004-A.  
XX  
XX 15-JAN-2000.  
XX  
XX 25-JUN-1998; 98KR-0024064.  
XX

PR 25-JUN-1998; 98KR-0024064.  
XX  
XX (SMSU ) SAMSUNG MEDICAL CENT.  
PA (JIND/) JIN D G.  
XX  
XX Jin DG;  
XX  
XX WPI; 2000-662424/64.  
XX  
XX Plasmid vector arraying neuropsychiatric gene more than copy number  
XX containing trinucleotide repeats its use thereof -  
XX  
XX Disclosure; Page 12; 23pp; Korean.  
XX  
XX This invention relates to the use of a plasmid vector array  
XX containing a neuropsychiatric gene containing trinucleotide repeats.  
XX Tri-nucleotide repeats and their expansion are known to be involved  
XX in a number of diseases including fragile X syndrome, spinocerebellar  
XX ataxia type III, hereditary hypertrophic, cardiomyopathy, dentatorubral  
XX and pallidolysian atrophy (DRPLA), spinocerebellar ataxia, Marfan  
XX syndrome; X-linked spinobular atrophy. The present sequence  
XX represents a nucleotide sequence of the invention showing the  
XX the CAG tri-nucleotide repeats.  
XX  
XX SQ Sequence 277 BP; 80 A; 66 C; 62 G; 69 T; 0 other;  
Query Match 4.3%; Score 53.4; DB 21; Length 277;  
Best Local Similarity 59.6%; Pred. No. 0.00014;  
Matches 90; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
Qy 873 GCTGCTGCTGCTGCGCGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932  
Db 215 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156  
Qy 933 GCTGTAGAAGAAAAAGAGGATTCGTATTCAATTTCAAAAGAAATCTGAAAAAGAGAGA 992  
Db 155 GCTGTTGCTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96  
Qy 993 CAAACAAAGAACTGAGACAGAGAAAGTGAAA 1023  
Db 95 AACTTAAAGAAATAAATACACCATGAGAAA 65  
RESULT 12  
ABL29756  
ID ABL29756 standard; DNA; 7029 BP.  
XX AC ABL29756;  
XX  
XX  
XX DT 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 40741.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:41:25 ; Search time 2972.12 Seconds  
(without alignments)  
10140.072 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240

Sequence: 1 aggtgtgagtcagccaaca.....gtaatacaactgtagtatag 1240

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estov:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	786.6	63.4	1201	13	BX399881
2	544	43.9	941	13	BX403420
3	510.8	41.2	917	13	BX370118
4	496.6	40.0	1201	13	BX358190

5	457.6	36.9	970	13	BX370209
6	328.2	26.5	1510	11	AK086973
7	308.2	24.9	784	12	BG863468
8	278.6	22.5	794	10	BF784177
9	268	21.6	1148	13	BX370210
10	265.8	21.4	725	10	BG740428
11	259	20.9	631	10	BG663870
12	245.2	19.9	527	10	BE032610
13	224.6	18.1	1201	13	BX358189
14	208.6	16.8	556	4	BX511978
15	197.4	15.9	517	10	BE476432
16	190.8	15.4	622	29	AG069679
17	190.8	15.4	677	29	AG107877
18	190	15.3	524	9	AV601192
19	185.6	15.0	471	9	AA265274
20	185.6	15.0	611	10	AW990468
21	185.6	15.0	691	14	BY708046
22	185.6	15.0	931	10	BG173684
23	185.6	15.0	1552	11	AK008060
24	185.6	15.0	1552	11	AK078934
25	184.6	14.9	486	9	AI425363
26	184	14.8	663	10	BE376590
27	182.4	14.7	650	10	BB625217
28	168.2	13.6	600	28	AZ379623
29	156.2	12.6	388	14	CB805443
30	154.8	12.5	216	10	BE206666
31	152.8	12.3	556	9	AI788300
32	151.2	12.2	530	9	AI790785
33	145.4	11.7	423	10	BB846577
34	143.8	11.6	291	28	AZ411779
35	133.6	10.8	426	13	BY033616
36	130.2	10.5	351	13	BY100408
37	128.8	10.4	471	10	BF040046
38	127	10.2	394	10	BB846133
39	124	10.0	685	29	AG142221
40	122.2	9.9	346	13	BY143304
41	121.2	9.8	369	13	BY036814
42	110	8.9	785	12	BI452873
43	107.6	8.7	1201	13	BX396896
44	102.8	8.3	889	12	BI454276
45	100	8.1	185	14	N47851

#### ALIGNMENTS

RESULT 1  
BX399881  
LOCUS  
DEFINITION  
ACCSSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BX399881 1201 bp mRNA linear EST 13-MAY-2003  
Clone CSOD1085YA24 5-PRIME, mRNA sequence.

BX399881  
EST.  
BX399881.1 GI:30622019

EST.  
Homo sapiens (human)

Homo sapiens

Eukaryote, Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10299.f, and

it belongs to a clone representative of this cluster. For more

information about this cluster and the virtual cDNA, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSOD1085BA12QPI&cluster=10299.f. Contact :

Peng Liang Email : fliang@lifetech.com URL :

BX370209 BX370209  
AK086973 Mus muscu  
BG863468 602796941  
BF784177 602108039  
BX370210 BX370210  
BG740428 602633817  
BG663870 BB663870  
BE032610 132035 MA  
BX358189 BX358189  
BX511978 RZPD Mus  
BA476432 159670 BA  
AG069679 Pan trogl  
AG107877 Pan trogl  
AV601192 AV601192  
AA265274 mx91d09.r  
AW990468 uf37d05.y  
BY708046 BY708046  
BG173684 602336823  
AK008060 Mus muscu  
AK078934 Mus muscu  
AI425363 mx91d09.y  
BE376590 601226419  
BB625217 BB625217  
AZ379623 1M0134G22  
CB805443 AMGNUC:C  
BG206666 RST26117  
AI788300 uk56d01.y  
AI790785 uk28a12.y  
BB846577 BB846577  
AZ411779 1M0185B04  
BY033616 BY033616  
BY100408 BY100408  
BF040046 BP250004B  
BB846133 BB846133  
AG142221 Pan trogl  
BY143304 BY143304  
BY036814 BY036814  
BI452873 603170211  
BX396896 BX396896  
BI454276 603171509  
N47851 yw95h05.r1

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0D1085BA12QPI.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1085YA24"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="First strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 290 a 285 c 301 g 297 t 28 others  
ORIGIN  
Query Match 63.4%; Score 786.6; DB 13; Length 1201;  
Best Local Similarity 93.3%; Pred. No. 8.7e-133;  
Matches 839; Conservative 21; Mismatches 35; Indels 4; Gaps 4;  
Qy 115 AGGTTCTGGTCTGTAATGAGTATAGAGAGGCCCCCAGAGATGCAACAGTCTTGAAGG 174  
Db 270 AAGTTCTGGTCTGTAATGAGTATAGAGAGGCCCCCAGAGATGCAACAGTCTTGAAGG 329  
Qy 175 CTCCAGGCTCGCTTCAACTGACCGTCTCCAGGCTGGAAGTCAATCATGTGGGCTCT 234  
Db 330 CTCCAGGCTCGCTTCAACTGACCGTCTCCAGGCTGGAAGTCAATCATGTGGGCTCT 389  
Qy 235 CAGTGACATGTTGTTGTAAGCGTCAAGCCCATGAGCCCATCATCAATGACCCGCTT 294  
Db 390 CAGTGACATGTTGTTGTAAGCGTCAAGCCCATGAGCCCATCATCAATGACCCGCTT 449  
Qy 295 CACCTCTCAGAGTACGACGAGGCGGGAATCTTCACTCGAGATGATCATCCCAATGT 354  
Db 450 CACCTCTCAGAGTACGACGAGGCGGGAATCTTCACTCGAGATGATCATCCCAATGT 509  
Qy 355 GGAGCCAGTATTCGGGGAACATCAGATGAGCTCCAGAACAGTCCCTGCGATGATC 414  
Db 510 GGAGCCAGTATTCGGGGAACATCAGATGAGCTCCAGAACAGTCCCTGCGATGATC 569  
Qy 415 TGCTTACCTTACCTGCAAGTTATGGGAGAGCTGTTCATTCCTCAGTGTATATCTTGTAGT 474  
Db 570 TGCTTACCTTACCTGCAAGTTATGGGAGAGCTGTTCATTCCTCAGTGTATATCTTGTAGT 629  
Qy 475 CGCTGAGAAAGAACTTGTGAAGTACTTGTCTACCTCAGTGGAGCCCGCTCCCGGA 534  
Db 630 CGCTGAGAAAGAACTTGTGAAGTACTTGTCTACCTCAGTGGAGCCCGCTCCCGGA 689  
Qy 535 TATTTCTGGAGCTCGGTCTCTCTGCTCAGCCATTCAGCTATTTATTTGTTCCGGAGCC 594  
Db 690 TATTTCTGGAGCTCGGTCTCTCTGCTCAGCCATTCAGCTATTTATTTGTTCCGGAGCC 749  
Qy 595 CAGGACCTTCAAGTGCAGTACGATCTTGGTCTGACCCCAAGAGCAATGGGACTTT 654  
Db 750 CAGGACCTTCAAGTGCAGTACGATCTTGGTCTGACCCCAAGAGCAATGGGACTTT 809  
Qy 655 GACTTGGTGCTACCTGGAAGAGCCTGAAGCCCGCAAGTCTGCACTGTAATCTCAC 714  
Db 810 GACTTGGTGCTACCTGGAAGAGCCTGAAGCCCGCAAGTCTGCACTGTAATCTCAC 869  
Qy 715 TGTGATTCGGTGTCCCAAGACACTGGAGTGTGTTAATATTCAGGTGTTATATCAAG 774  
Db 870 TGTGATTCGGTGTCCCAAGACACTGGAGTGTGTTAATATTCAGGTGTTATATCAAG 928  
Qy 775 TTTTACCGAGTTAGGTTTTTCAATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCCAG 834  
Db 929 -TTTACCGAGTTAGGTTTTTCAATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCCAG 987  
Qy 835 CACCATGCTTCTGACGCGGAGCTGTACTTCTTACAAATAGCTGTGCTGCGCCGCTCG 894  
Db 988 SACCATGCTTCTGACGCGGAGTGTAATCTATATAATATGCTGTGCTGCGCCGCTCG 1047

Qy 895 TTGTTGTGGTGCACACTGCTGCTCCGTTGTTGTTTCTGCTGTAGAGAAAAGAGGATT 954  
Db 1048 TTGATGTGCTGCAACT-CTGCKRCGGTANTTGTCTTCTGTAGAA-RAAARARGGTT 1105  
Qy 955 TCCTATTCAATTTCAAGAGAAATCTGAAAAGAGAGAGCAACAAGAACTGACAC 1013  
Db 1106 YSGTATCMAATTTCAARAATCTGAAAAGAGAGAGCAACAAGAACTGACAC 1164  
RESULT 2  
BX403420 941 bp mRNA linear EST 13-MAY-2003  
LOCUS BX403420 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS0D1085YA24 5-PRIME, mRNA sequence.  
ACCESSION BX403420  
VERSION BX403420.1 GI:30607302  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 941)  
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
10299.f, and it belongs to a clone representative of this cluster.  
For more information about this cluster and the virtual cDNA, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS5AA017ZE02RM1&cluster=10299.f. Contact :  
Peng Liang Email : fliang@lifetech.com URL : Corporation 1600  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS5AA017ZE02RM1.  
Location/Qualifiers  
1. .941  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1085YA24"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="First strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 214 a 241 c 255 g 227 t 4 others  
ORIGIN  
Query Match 43.9%; Score 544; DB 13; Length 941;  
Best Local Similarity 98.8%; Pred. No. 8.5e-89;  
Matches 569; Conservative 0; Mismatches 5; Indels 2; Gaps 2;  
Qy 115 AGGTTCTGGTCTGTAATGAGTATAGAGGCCCCCAGAGTCAACAGTCTTGAAGG 174  
Db 284 AAGTTCTGGTCTGTAATGAGTATAGAGGCCCCCAGAGTCAACAGTCTTGAAGG 343  
Qy 175 CTCCAGGCTCGCTTCAACTGACCGTCTCCAGGCTGGAAGCTCATGTGGGCTCT 234  
Db 344 CTCCAGGCTCGCTTCAACTGACCGTCTCCAGGCTGGAAGCTCATGTGGGCTCT 403  
Qy 235 CAGTGACATGTTGTTGTAAGCGTCAAGCCCATGAGCCCATCATCAATGACCCGCTT 294  
Db 404 CAGTGACATGTTGTTGTAAGCGTCAAGCCCATGAGCCCATCATCAATGACCCGCTT 463  
Qy 295 CACCTCTCAGAGTACGACGAGGCGGGAATCTTCACTCGAGATGATCATCCCAATGT 354  
Db 464 CACCTCTCAGAGTACGACGAGGCGGGAATCTTCACTCGAGATGATCATCCCAATGT 523

```
QY 355 GGAGCCAGTGATTGGGGAACATCAGATGAGCTCCAGACAGTGCCTGCGATGATC 414
Db 524 GGAGCCAGTGATTGGGGAACATCAGATGAGCTCCAGACAGTGCCTGCGATGATC 583
QY 415 TGCTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCACTGTTAATCTTGTAGT 474
Db 584 TGCTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCAATCCCACTGTTAATCTTGTAGT 643
QY 475 CGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTTCACTGGAACCCGGCTCCCGGA 534
Db 644 CGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTTCACTGGAACCCGGCTCCCGGA 703
QY 535 TATTTCCTGGAGCTCGGTCTCTGTCAGCCATTCAGAGCTATTATTTG-TTCGGGAGC 593
Db 704 TATTTCCTGGAGCTCGGTCTCTGTCAGCCATTCAGAGCTATTATTTGTTCCGGAGC 763
QY 594 CCACGACCTTCAAAGTGCAGTGAGCATCTCTGCTTGCACCCACAGACGCAATGGGACTT 653
Db 764 CCACGACCTTCAAAGTGCAGTGAGCATCTCTGCTTGCACCC-CAGAGCAATGGGACTT 822
QY 654 TGACTTCGGTGGCTACCTGGAGAGCTTGAAGGCC 689
Db 823 TGACTTCGGTGGCTTCCCTGGAAGAGCTTGAAGGCC 858

RESULT 3
LOCUS BX370118 917 bp mRNA linear EST 08-MAY-2003
DEFINITION BX370118 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1076YN22 5-PRIME, mRNA sequence.
ACCESSION BX370118
VERSION BX370118.1 GI:30447910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2250.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG037ZE05_CS03493_2&cluster=2250.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG037ZE05_CS03493_2.
FEATURES
source
1. .917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1076YN22"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoK V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 272 a 203 c 190 g 252 t
ORIGIN
Query Match 41.2%; Score 510.8; DB 13; Length 917;
Best Local Similarity 99.2%; Pred.No. 8.8e-83;
Matches 524; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 713 ACTGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTATATATCCAGGTGATTATCA 772
```

```
Db 9 ACTGTGATTCGGTGTCCCA--GAACCTGGAGGTGGTATTATATATTCAGGTGATTATCA 66
QY 773 AGTTTACCGAGTTTAGGTTTTTCTACTTGGGGCAAAAGTTGGACTTGGACTAGCA 832
Db 67 AGTTTACCGAGTTTAGGTTTTTCTACTTGGGGCAAAAGTTGGACTTGGACTAGCA 126
QY 833 GGCAACCATGCTTCTGAAGCCGACGCTGTAATCTTACAATACGCTGCTGCTGCCCGCT 892
Db 127 GGCAACCATGCTTCTGAAGCCGACGCTGTAATCTTACAATACGCTGCTGCTGCCCGCT 186
QY 893 CGTTGTTGTCGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
Db 187 CGTTGTTGTCGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
QY 953 TTTTCGTTATTCATTTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACA 1012
Db 247 TTTTCGTTATTCATTTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACA 306
QY 1013 GAAAGTGAATGAAATCCCGCTACAAATTCAGATGAACAAAGACACAGAAACCGCT 1072
Db 307 GAAAGTGAATGAAATCCCGCTACAAATTCAGATGAACAAAGACACAGAAACCGCT 366
QY 1073 TCTCTCCCTCCCAATCTGTGAATCCAGTGATCTGAAAGAGAGAGAGAGAGAGAGAGAGAG 1132
Db 367 TCTCTCCCTCCCAATCTGTGAATCCAGTGATCTGAAAGAGAGAGAGAGAGAGAGAGAGAG 426
QY 1133 CCTCTCACGAGCGGCTGATCAACGCTCCACCCAGGCGAGCAAGTCATCCACAGCTTCT 1192
Db 427 CCTCTCACGAGCGGCTGATCAACGCTCCACCCAGGCGAGCAAGTCATCCACAGCTTCT 486
QY 1193 TTTAATCTGGCCAGTCTGAGAGGTGAGTAAATCACTAGTATAGTAGTAGTAGTAGTAGTAG 1240
Db 487 TTTAATCTGGCCAGTCTGAGAGGTGAGTAAATCACTAGTATAGTAGTAGTAGTAGTAGTAG 534

RESULT 4
LOCUS BX358190 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION BX358190 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1035YC01 5-PRIME, mRNA sequence.
ACCESSION BX358190
VERSION BX358190.1 GI:30372233
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1035AB01QF1.
FEATURES
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1035YC01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoK V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

BASE COUNT	341 a	245 c	245 g	319 t	51 others
ORIGIN					
Query Match	40.0%;	Score 496.6;	DB 13;	Length 1201;	
Best Local Similarity	95.5%;	Pred. No. 3.2e-80;			
Matches 511;	Conservative	0;	Mismatches 24;	Indels 0;	Gaps 0;
QY	706	AAATCTCAGTGTGATTCGGTGTGCCCCAAGACACTGGAGTGTGTTATTAATTCAGGTGT	765		
Db	27	AAAAAGCAGCTGTGATCCGGTCCGGAATTCGGGATGTGTGTTATTAATTCAGGTGT	86		
QY	766	ATTATCAAGTTACCGAGTTTAGTGTTCATTCCTACTTGGGCAAGTTGACATTGG	825		
Db	87	ATTATCAAGTTACCGAGTTTAGTGTTCATTCCTACTTGGGCAAGTTGACATTGG	146		
QY	826	ACTAGCAGGCACCATCTCTGACGCCGAGTGTACTTCAATACACCTGCTGCTGCTG	885		
Db	147	ACTAGCAGGCACCATCTCTGACGCCGAGTGTACTTCAATACACCTGCTGCTGCTG	206		
QY	886	CCGCGTCTGTTGTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	945		
Db	207	CCGCGTCTGTTGTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	266		
QY	946	AAGAGGATTTCTGATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAAAC	1005		
Db	267	AAGAGGATTTCTGATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAAAC	326		
QY	1006	TGAGACAGAAAGTGAATGAACCTCCGGCTACAAATCAGATGAACAAAGACCCACAGA	1065		
Db	327	TGAGACAGAAAGTGAATGAACCTCCGGCTACAAATCAGATGAACAAAGACCCACAGA	386		
QY	1066	AACCGCTTCTCTCCCTCCCAATCTCTGTAATCCAGTATCTGTAACAAAGAAACAGTAG	1125		
Db	387	AACCGCTTCTCTCCCTCCCAATCTCTGTAATCCAGTATCTGTAACAAAGAAACAGTAG	446		
QY	1126	CTGTGGCCCTCTCACAGGGCTGATCAAGTCCACCGCCAGCAGTCTCATCCACA	1185		
Db	447	CTGTGGCCCTCTCACAGGGCTGATCAAGTCCACCGCCAGCAGTCTCATCCACA	506		
QY	1186	GGCTTCTTTTAAATCTGCGCAGTCTCAGAGAGTTCAGTATACAACTGTAGTATAG	1240		
Db	507	GGCTTCTTTTAAATCTGCGCAGTCTCAGAGAGTTCAGTATACAACTGTAGTATAG	561		
RESULT 5					
LOCUS	BX370209	970 bp	mRNA	linear	EST 08-MAY-2003
DEFINITION	BX370209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA				
ACCESSION	BX370209				
VERSION	BX370209.1	GI:30451880			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Li, W.B., Gruber, C.; Jesse, J., and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
	Library was constructed by Life Technologies, a division of				
	Invitrogen. This sequence belongs to sequence cluster 10299.f, and				
	it belongs to a clone representative of this cluster. For more				
	information about this cluster and the virtual cDNA, see				
	http://www.genoscope.cns.fr/				
	cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_CS03685_1cluster=10299.f.				
	Contact : Feng Liang Email : fliang@lifetech.com URL :				
	http://fulllength.invitrogen.com/ invitrogen Corporation 1600				
	Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05_CS03685_1.				



774 GTTACCGAGTTAGTTTTCATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCAG 833  
 799 -----AGGCCGACACTGCCGACCTGGGCCCATCATCTGCTGCGAGTGGCCT 845  
 834 GCACCATGCTTCTGACGCGAGCTGACTCTTACAAATACGCTGCTGCTGCGCGCGTC 893  
 846 TTTCCTTGTCTGTGATCTGATCAITGTTTGAATATAATATTCGTGCTGTTGT--- 901  
 894 GTTGTGTGCTGCAACTGCTGCTGCGGTTGTTGTTTCTGCTGTGAGAGAAAAGAGGAT 953  
 902 -----GCCTCCAGGAGAGAAAAGGAAGAAATCTACTTATCAAAATGAAATAGGAAA 952  
 954 TTCTGATTCAAATTCAGAAAGAAATCTGAAAGAGAGAGACAAACAAAGAAATGAGACAG 1013  
 953 TCTGCAACATGAGGACAA-----ACAAAGCAGATCCGGAGACAAAGTTAA 998  
 1014 AAGTGGAAATGAAACTCCGCTACAAATCAGATGAAACAAAGACCCAGAAACCGCTT 1073  
 999 AAGTGGAAAGAAAACACTACGGGTACAGTTCGGATGAGGCAAGGCTGCACAGACTGCAT 1058  
 1074 CTCTCCCTCCAAATCCTGTGAATCCAGTGATCCTGAAACAAAGAAACAGTAGC 1126  
 1059 CTCTCCCTCCTAAATCTGCTGAAGTCAGCTTCCAGAAACACGACGAGTAGC 1111

RESULT 7  
 BG863468  
 LOCUS 602796941F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4917803 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG863468  
 VERSION BG863468.1 GI:14214006  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth  
 Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHAM0830 row: b column: 12  
 High quality sequence stop: 768.  
 Location/Qualifiers  
 1..784  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NMRI"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4917803"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Mam4"  
 /notes="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Priscilla Furth,  
 NIH Reference for transgenic model: Li et al., Cell Growth  
 and Differentiation 7, 3-11 (1996)."  
 188 a 202 c 208 g 186 t

BASE COUNT  
 ORIGIN

Query Match 24.9%; Score 308.2; DB 12; Length 784;  
 Best Local Similarity 70.8%; Pred. No. 4.8e-46;  
 Matches 437; Conservative 0; Mismatches 178; Indels 2; Gaps 2;  
 115 AGTTCTGGGCTCTGTAATGAAGTATAGAGAGCCGCCAGAGCTCATCATGTGGGCTCT 234  
 149 AGCTTCGGATCCAGTTATCAGATCATAGAGGCTCTCAGAGTGTAAACAGTCTCTAAGGA 208  
 175 CTCCAGGCTCGCTTCAACTGACACCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCT 234  
 209 CTGAGAGGCTCACTTCAACTGCAACCTGACTCAGGCTGGAAGCTTCTCATGTGGACTCT 268  
 235 CAGTGCATGCTGCTGTAAGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTC 294  
 269 TAACCAATGCTGCTGCTGAGTCTCACCACCAAGGACCATCATCAACCAACACCGCTT 328  
 295 CACTCTCAGAGTACAGCAGGCGGGAATCTCCTGGAGATGATCATCCACATCT 354  
 329 CACTATGCCAGTTTAAACAGCACTGACAGCTTCACTCGGAGTTGATCATCTCATGATGT 388  
 355 GGAGCCAGTGATTCGGGGAAACATCAGATGCGAGCTCCAGAAACAGTCCGCTGCATGGATC 414  
 389 GCAGCCAGTGAATCGGGATCCGTCGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 448  
 415 TGCTTACCTTACCGTCCAGTTATGGAGAGCTGTTCTATCCAGTGTAAATCTTGTAGT 474  
 449 TGCCTTCTCTCAGTCAAGTCAATGGGACCTTCAATTCCTAGCAACACCTTATAGT 508  
 475 CGCTGAGAAATGAACCTTGTGAAGTCTTCTGCTTACCTCAGCTGAGACCGGCTCCCGGA 534  
 509 CACTGAGGTTGAACCTTGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 568  
 535 TATTTCTGGAGCTCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594  
 569 TATTTCTGGAGCTTGAAGTCTCCGCTGAGCTTCAAGTTACAAATCTCTTCTTGGAGCT 628  
 595 CAGGAGCTTCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 654  
 629 GGGCAACTTTATAGAGGCTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAG 687  
 655 GACTTGGCTGGCTACCTTGAAGAGCTTGAAGAGCGGCTGGAAGCTGCAACTGTAAATCTCAC 714  
 688 GACTTGTGTCGAGAGCTGAGGACTTGCAGGCGCAGGAGCTTAACTGTCTCAA-CTGAC 746  
 715 TGTGATTCGGTGTCCCT 731  
 747 TGTGATTCAGCTTCCAC 763

BF784177 794 bp mRNA linear EST 12-JAN-2001  
 602108039F1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:4236489  
 5', mRNA sequence.  
 BF784177  
 BF784177.1 GI:12089213  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: L1AM9844 row: n column: 10  
High quality sequence stop: 694.  
Location/Qualifiers

## FEATURES

source

1. .794  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4236489"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI CGAP Kid14"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. |"  
BASE COUNT 186 a 214 c 215 g 179 t  
ORIGIN  
Query Match 22.5%; Score 278.6; DB 10; Length 794;  
Best Local Similarity 69.6%; Pred. No. 1.1e-40;  
Matches 406; Conservative 0; Mismatches 174; Indels 3; Gaps 2;  
QY 115 AGGTTCTGGGTCTGGTAATGAAGTATAGAGAGGCCCCAGATGCAAGTCCTGGAAGGG 174  
Db 149 AGCTTCGGATCCAGTTATCAGATCATAGAGGTCCTCAGATGTACAGTCTTAAGGA 208  
QY 175 CTCCAGGCTGCTTCAACTGACCGTCTCCAGGGTGAAGTCAATCATGTGGGCTCT 234  
Db 209 CTCAGAGGCTCACTTCAACTGCACCGTGAAGTCTCATGTGGAAGTCT 268  
QY 235 CAGTGACATGCTGTGTAAGCTGAGCCATGAGCCATCATCACCATGACCGCTT 294  
Db 269 TAACCAATGGTGTGCTGAGTCTACACCCAGGACCCATCATCACCACCAACCGCTT 328  
QY 295 CACCTCTCAGAGTACGACGAGGCGGGAATTCACCTCGGAGATGATCATCAACAATGT 354  
Db 329 CACCTATGCCAGTTACAAACGACCTGACAGCTTCATCTCGGAGTTGATCATCATGATGT 388  
QY 355 GGAGCCAGGATTCGGGGACATCAGATGAGCTCCAGACAGTGCCTGCAAGATC 414  
Db 389 GCAGCCAGTGAATCGGGATCCGTCGAATGAGCTGAGACAGCCATGAGGATGTCGATC 448  
QY 415 TGCTTACCTTACCGTCAAGTATGAGGAGAGCTGTTTCAATCCAGCTGTAATCTGTAGT 474  
Db 449 TGCTTCTCTCAGTGAATGATGAGGACCTGAACTTCCTAGCAACACCTTATAGT 508  
QY 475 CGCTGAGATGAACCTTGTGAAGTACTTGTACTACCTCACTGAGCCCGGCTCCCGGA 534  
Db 509 CACTGAGGTTGAACCTGTGAATGTGACTTGTATGCGGTGGGCTGAGACCTCACTCCCGGA 568  
QY 535 TATTTCTGGAGCTGGTCTCTGCTGAGCCATTCAGCTATATATTTTGTTCGGAGCC 594  
Db 569 TATTTCTGGAGCTGAGGTTCCCGTAAGCCATTGAGTTACAATTCCTTTCTGGAGCC 628  
QY 595 CAGCGACCTTCAAGTGCAGTGAAGTCTGAGTCTGAGCCCTGACCCCAAGAGCAATGGGCTTT 654  
Db 629 --GGGCACTTTAGAGGCTGAGTCTGCTGGAAGTCA--CACACTGGGCAAGGACCTT 685  
QY 655 GACTTGGTGGCTGACTGGAAGAGCTGGAAGGCGCGCAAGTCT 697  
Db 686 GACTTGTGTGCAAGAGCTGAAGGACTTGCAGGCCAGCAGTCT 728

## RESULT 9

BX370210

LOCUS

DEFINITION

BX370210 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CS0D1085YA24 5-PRIME, mRNA sequence.

ACCESSION

BX370210

VERSION

BX370210.1 GI:30451881

KEYWORDS

EST.

BX370210 1148 bp mRNA linear EST 08-MAY-2003

BX370210 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CS0D1085YA24 5-PRIME, mRNA sequence.

ACCESSION

BX370210

VERSION

BX370210.1 GI:30451881

KEYWORDS

EST.

## SOURCE

ORGANISM

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 1148)

AUTHORS

Li W.B., Gruber C., Jessee, J. and Polayes, D.

JOURNAL

Full-length cDNA libraries and normalization

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10299.f, and

it belongs to a clone representative of this cluster. For more

information about this cluster and the virtual cDNA, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAG039ZE05\_CS03685\_2&amp;cluster=10299.f.

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05\_CS03685\_2.

Location/Qualifiers

1. .1148

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1085YA24"

/tissue types="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 204 a 263 c 137 g 544 t

ORIGIN

Query Match 21.6%; Score 268; DB 13; Length 1148;

Best Local Similarity 89.4%; Pred. No. 8.9e-39;

Matches 322; Conservative 0; Mismatches 35; Indels 3; Gaps 3;

QY 716 GTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTATATATCCAGGTGATTATCAAGT 775

Db 1 GTGATTCGGTGT-TCCAGACACTGGAGGTGGTATTATATATCCAGGTGATTATCAAGT 59

QY 776 TTACCGAGTTAGGTTTTTCAATTCCTTACCTTGGGGCAAGTTGGACTTGGACTAGCAGGC 835

Db 60 TTACCGAGTTAGGTTTTTCAATTCCTTACCTTGGGGCAAGTTGGACTTGGACTAGCAGGC 119

QY 836 ACCATGCTTCTGACGCGGAGCTGCTTACATACGCTGCTGCTGCGCCGCTCGT 895

Db 120 ACCATGCTTCTGACGCGGAGCTGCTTACATACGCTGCTGCTGCGCCGCTCGT 179

QY 896 TGTGTGGCTGCAACTGCTGCTGCGGTTGTTTCTGCTGTAGAGAAAAGAGGATTT 955

Db 180 TGTGTGGCTGCAACTGCTGCTGCGGTTGTTTCTGCTGTAGAGAAAAGAGGATTT 239

QY 956 CGTATTCAATTTCAAAGAAA-TCTGAAAAGAGAGACAAACAAAGAACTGACAGAGA 1014

Db 240 TGTATTATATTTAAAGAAATTTTGTAAAGTGAAGACACTCAGTGAATCTGACAGAGA 299

QY 1015 AAG-TGGAATGAAGAACTCGGCTACAAATTCAGATGAACAAAGACACAGAACCGCTT 1073

Db 300 AAGTTGTTAATGAAGAACTCGGCTTTCATTTCTTTTGTGACCAAGTAATCCGTT 359

## RESULT 10

BG740428

LOCUS

DEFINITION

BG740428 Homo sapiens cDNA clone IMAGE:477895

mRNA sequence.

ACCESSION

BG740428

VERSION

BG740428.1 GI:14051081

KEYWORDS

EST.







LOCUS BE476432 517 bp mRNA linear EST 27-MAR-2003  
DEFINITION L59670 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION BE476432  
VERSION BE476432.1 GI:9595965  
KEYWORDS EST.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 517)  
Cho.J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D. and  
Sonstegard,T., Capuco,A.V., White,J., Van Tassell,C.P., Connor,E.B.  
Quackenbush,J.  
TITLE Analysis of bovine mammary gland EST and functional annotation of  
the Bos taurus gene index  
JOURNAL Mamm. Genome 13 (7), 373-379 (2002)  
MEDLINE 22135956  
PUBMED 12140684  
COMMENT Contact: Sonstegard TS  
USDA, ARS, Beltsville Agricultural Research Center  
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA  
Tel: 301 504 8416  
Fax: 301 504 8414  
Email: tads@psi.barc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCACGTCAGCAGC  
Plate: 27 row: H column: 19  
Seq primer: ATTAGTGACACTATAG.  
Location/Qualifiers  
1..517  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="BARC 5BOV"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled mRNA isolated from mammary  
tissues at eight physiological, developmental, and disease  
states."

FEATURES  
source

BASE COUNT 117 a 133 c 165 g 102 t  
ORIGIN  
Query Match 15.9%; Score 197.4; DB 10; Length 517;  
Best Local Similarity 76.2%; Pred. No. 6.1e-26;  
Matches 243; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
QY 113 GTAGTTCTGGTCTGGTAAATGAAGTCATAGAGCCGCCCAAGATGCAAGTCTCTGAAG 172  
Db 199 GCAGCCTGTGGATCCAGCAGTGAATCATAGAGGGTCCCAAGATGTACAGCCCTGAAG 258  
QY 173 GGCTCCCAAGGTCGGTTCACATGACCGTCTCCCAAGGCTGGAAGTCATCATGTGGGCT 232  
Db 259 GGCTCGGAGGCTCGTTCACTGACCACTCTCCAGGGCTGGAAGTCGTATGTGGGCT 318  
QY 233 CTCAGTGATCGTGGTCTAAGCGTCAAGCGCCCATGAGCCCATCATCAATGACCGC 292  
Db 319 CTGAGAGGCACAGTGGTCTGAGCATGACACCTAATGAGACCATCATCACCAGTGACCGC 378  
QY 293 TTCACCTCTCAGAGTACACAGCGGGGAACTTCACCTCGGAGATGATCATCCACAAT 352  
Db 379 TTCCTTCGGCAAGCTACCAAGAGGGCGGAACTTCATCTGTGATGATAATTCATGAC 438  
QY 353 GTGAGAGCCAGTGATTGGGGAAACATCAGATGAGCGCTCCAGACAGTCGCTGCATGA 412  
Db 439 GTGCNACTGAGGATCGCGGCAAGTCAATGACGCTTCAGAACAGCAATCGGGATGA 498  
QY 413 TCTGTTACCTTACCGTCC 431

Db 499 GCCGCCTTCCTTTCTGTC 517  
Search completed: November 12, 2003, 23:19:39  
Job time : 2974.12 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:42:00 ; Search time 82.0207 Seconds  
(without alignments)  
6672.894 Million cell updates/sec

Title: US-09-729-264-5  
Perfect score: 1240  
Sequence: 1 aggtgtgagtcagcccaaca.....gtaatacaactgtatag 1240

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCUS COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	51.8	4.2	397	3	US-09-253-691-3
C 2	48.2	3.9	325	2	US-08-531-927B-3
C 3	47.6	3.8	2733	4	US-08-997-685A-1
C 4	44.8	3.6	253	4	US-09-491-356C-13
C 5	44.8	3.6	285	4	US-09-491-356C-14
C 6	44.8	3.6	265	4	US-09-491-356C-16
C 7	44.8	3.6	265	4	US-09-491-356C-17
C 8	44.8	3.6	265	4	US-09-491-356C-18
C 9	44.8	3.6	6794	4	US-09-491-356C-2
C 10	44.8	3.6	55298	4	US-09-491-356C-1
C 11	44.2	3.6	319608	4	US-09-539-333D-1
C 12	44.2	3.6	319608	4	US-09-679-409-1
C 13	43.4	3.5	3302	4	US-09-620-312D-475
C 14	43.2	3.5	2294	4	US-09-086-663A-70
C 15	43.2	3.5	3334	4	US-09-086-663A-7
C 16	43	3.5	6558	4	US-09-491-356C-7
C 17	42.8	3.5	1776	2	US-08-531-927B-1
C 18	42.8	3.5	1776	3	US-09-041-886-12
C 19	42.6	3.4	1037	4	US-09-181-585-3
C 20	42.6	3.4	1159	4	US-09-181-585-1
C 21	42.6	3.4	1471	4	US-09-181-585-2
C 22	42.6	3.4	3292	1	US-07-814-964-12
C 23	42.6	3.4	3292	1	US-08-258-442-12
C 24	42.6	3.4	3292	1	US-08-328-809-7
C 25	42.6	3.4	3292	4	US-08-866-840-7
C 26	42.6	3.4	3292	5	FCT-US92-11107-12
C 27	42.6	3.4	40000	4	US-09-780-049-18

C 28	42	3.4	2674	4	US-09-817-180-1	Sequence 1, Appli
C 29	41.6	3.4	78	3	US-09-043-303-12	Sequence 12, Appli
C 30	41.6	3.4	203	3	US-09-043-303-7	Sequence 7, Appli
C 31	41.6	3.4	6835	4	US-09-125-635-1	Sequence 1, Appli
C 32	41.2	3.3	379	1	US-09-591-383-5	Sequence 5, Appli
C 33	41.2	3.3	379	1	US-08-145-617-5	Sequence 5, Appli
C 34	41.2	3.3	63588	4	US-09-873-404-3	Sequence 3, Appli
C 35	41.2	3.3	66804	4	US-09-740-041-3	Sequence 3, Appli
C 36	41	3.3	3771	1	US-08-185-432-3	Sequence 3, Appli
C 37	41	3.3	3771	1	US-08-185-432-23	Sequence 23, Appli
C 38	41	3.3	5063	1	US-08-185-432-1	Sequence 1, Appli
C 39	40.8	3.3	6177	4	US-08-479-913E-1	Sequence 1, Appli
C 40	40.8	3.3	7257	4	US-09-091-042A-1	Sequence 1, Appli
C 41	40.6	3.3	238	4	US-09-491-356C-15	Sequence 15, Appli
C 42	40.6	3.3	2156	4	US-09-086-663A-78	Sequence 78, Appli
C 43	40.6	3.3	4621	4	US-09-125-635-9	Sequence 9, Appli
C 44	40.4	3.3	484	6	549550-5	Patent No. 549550
C 45	40.4	3.3	10348	2	US-08-457-273B-41	Sequence 41, Appli

ALIGNMENTS

RESULT 1  
US-09-253-691-3/c  
; Sequence 3, Application US/09253691  
; Patent No. 6124100  
; GENERAL INFORMATION:  
; APPLICANT: Dong Kyu JIN  
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases  
; FILE REFERENCE: 1942/36  
; CURRENT APPLICATION NUMBER: US/09/253,691  
; CURRENT FILING DATE: 1999-02-22  
; EARLIER APPLICATION NUMBER: KR 98-6,278  
; EARLIER FILING DATE: 1996-02-26  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Wordperfect 6.1/Windows  
; SEQ ID NO 3  
; LENGTH: 397  
; TYPE: DNA  
; ORGANISM: human  
US-09-253-691-3

Query Match	4.2%	Score 51.8;	DB 3;	Length 397;
Best Local Similarity	58.9%	Pred. No. 3.8e-05;		
Matches	89;	Conservative	0;	Mismatches 62;
				Indels 0;
Gaps				0;
Qy	873	GCTGCTGCTGCTGCGCGCTGTTGTTGCTGCACTGCTGCTGCGTTGTTGTTGTTTCT	932	
Db	200	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	141	
Qy	933	GCTGTAGAGAAAAGAGGATTTCTGATTCATTTTCAAAAGAAATCTGAAAAGAGAGA	992	
Db	140	GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	81	
Qy	993	CAAAACAAGAACTGACAGACAGAAAGTGAAA	1023	
Db	80	AACTTAAAGAATAAATACCATGAGAAA	50	

RESULT 2  
US-08-531-927B-3/c  
; Sequence 3, Application US/08531927B  
; Patent No. 5840491  
; GENERAL INFORMATION:  
; APPLICANT: Kakizuka, Akira  
; TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph  
; Patent No. 5840491  
; TITLE OF INVENTION: Disease Gene and Uses Thereof  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.





Query Match 3.6%; Score 44.8; DB 4; Length 55298;

```

; LOCATION: 29386..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 64666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 65854..67854
; OTHER INFORMATION: 3'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 94124..94964
; OTHER INFORMATION: exon g35017
; FEATURE:
; NAME/KEY: exon
; LOCATION: 201188..201234
; OTHER INFORMATION: exon S g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 214676..214793
; OTHER INFORMATION: exon T g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215702..215746
; OTHER INFORMATION: exon U g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216836..216915
; OTHER INFORMATION: exon V g35030 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 213818..215818
; OTHER INFORMATION: 3'regulatory region g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215819..215941
; OTHER INFORMATION: exon R complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215819..215975
; OTHER INFORMATION: exon Rbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216661..216952
; OTHER INFORMATION: exon Qbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216661..217061
; OTHER INFORMATION: exon Q complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 217027..217061
; OTHER INFORMATION: exon Q1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 229647..229742
; OTHER INFORMATION: exon X complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 230408..230721
; OTHER INFORMATION: exon P complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231272..231412

```

```

; OTHER INFORMATION: exon Obis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231787..231880
; OTHER INFORMATION: exon O2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231870..231879
; OTHER INFORMATION: exon O1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 234174..234321
; OTHER INFORMATION: exon O complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 237406..237428
; OTHER INFORMATION: exon Nbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 239719..239807
; OTHER INFORMATION: exon N2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 239719..239853
; OTHER INFORMATION: exon N complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240569
; OTHER INFORMATION: exon M1117 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240596
; OTHER INFORMATION: exon M1090 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240617
; OTHER INFORMATION: exon M1069 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240644
; OTHER INFORMATION: exon MS2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240824
; OTHER INFORMATION: exon M862 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240994
; OTHER INFORMATION: exon M692 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..241685
; OTHER INFORMATION: exon M1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240800..240993
; OTHER INFORMATION: exon MS1 complement g34872 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 241686..243685
; OTHER INFORMATION: 5'regulatory region g34872 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 290652..292652
; OTHER INFORMATION: 3'regulatory region g34665 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 292653..292841

```

Query Match 3.6%; Score 44.2; DB 4; Length 319608;  
 Best Local Similarity 66.0%; Pred. No. 0.2;  
 Matches 64; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 24 TGGATCAGTCTTCTAGGCTGCATACAAACACACATTAACCTGCTGCTTAGAACATGG 83  
Db 31167 TGTATTAGTCTCTTGGCTGCATATAAAATAACACTGTCTGGTGGCTTAGAACACAT 31108  
QY 84 AAAGGCATTTGCTCAGGTTCCAGAGCTGTAGCTTC 120  
Db 31107 AAACCTATTCTCAGTTCTTAGGGCTGGAATCC 31071

RESULT 12  
US-09-679-409-1/c  
Sequence 1, Application US/09679409  
Patent No. 655316  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Essicux, Laurent  
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS  
FILE REFERENCE: 53 US15 CIP  
CURRENT APPLICATION NUMBER: US/09/679,409  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 09/539,333  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 09/416,384  
PRIOR FILING DATE: 1999-10-12  
PRIOR APPLICATION NUMBER: 60/168,088  
PRIOR FILING DATE: 1999-11-30  
NUMBER OF SEQ ID NOS: 134  
SOFTWARE: Patent.pm  
SEQ ID NO 1  
LENGTH: 319608  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 199122..201122  
OTHER INFORMATION: 5' regulatory region  
NAME/KEY: exon  
LOCATION: 201123..201234  
OTHER INFORMATION: exon S  
NAME/KEY: exon  
LOCATION: 201123..201560  
OTHER INFORMATION: exon S2  
NAME/KEY: exon  
LOCATION: 214676..214793  
OTHER INFORMATION: exon T  
NAME/KEY: exon  
LOCATION: 215702..215746  
OTHER INFORMATION: exon U  
NAME/KEY: exon  
LOCATION: 216836..216994  
OTHER INFORMATION: exon V  
NAME/KEY: exon  
LOCATION: 216836..217077  
OTHER INFORMATION: exon V2  
NAME/KEY: exon  
LOCATION: 217671..217764  
OTHER INFORMATION: exon V1  
NAME/KEY: exon  
LOCATION: 227655..227736  
OTHER INFORMATION: exon V4  
NAME/KEY: exon  
LOCATION: 238715..238919  
OTHER INFORMATION: exon V3  
NAME/KEY: exon  
LOCATION: 240440..240673  
OTHER INFORMATION: exon W  
NAME/KEY: exon  
LOCATION: 240440..241153  
OTHER INFORMATION: exon W2  
NAME/KEY: exon  
LOCATION: 205329

LOCATION: 241072..241291  
OTHER INFORMATION: exon X  
NAME/KEY: exon  
LOCATION: 244353..244561  
OTHER INFORMATION: exon Y  
NAME/KEY: exon  
LOCATION: 246273..247802  
OTHER INFORMATION: exon Z  
NAME/KEY: misc.feature  
LOCATION: 247803..249803  
OTHER INFORMATION: 3' regulatory region  
NAME/KEY: allele  
LOCATION: 8316  
OTHER INFORMATION: 99-27943-150 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 21672  
OTHER INFORMATION: 99-27935-193 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 65485  
OTHER INFORMATION: 8-128-33 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 95396  
OTHER INFORMATION: 99-31960-163 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 107281  
OTHER INFORMATION: 99-24656-260 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 160640  
OTHER INFORMATION: 99-24639-163 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 160876  
OTHER INFORMATION: 99-24634-108 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 168974  
OTHER INFORMATION: 99-7652-162 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 170810  
OTHER INFORMATION: 99-16100-147 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 173358  
OTHER INFORMATION: 99-5862-167 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 189957  
OTHER INFORMATION: 99-5919-215 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 197163  
OTHER INFORMATION: 99-24658-410 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 200778  
OTHER INFORMATION: 8-303-235 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 203651  
OTHER INFORMATION: 8-300-221 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 202679  
OTHER INFORMATION: 8-300-193 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 203378  
OTHER INFORMATION: 8-299-128 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 204138  
OTHER INFORMATION: 8-296-213 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 204605  
OTHER INFORMATION: 8-252-190 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 204934  
OTHER INFORMATION: 99-24644-194 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 205206  
OTHER INFORMATION: 8-295-248 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 205329

Qy	24	TGGATCAGTTTCTAGGCTGCATATACAAAGCACATACTGTGGCTTTAGAACATGG	83
Db	31167	TGTATTAGTCTCTTTGGCTGCATATATAAATAACACTGTCGTGGGTGGCTTTGAACAACAT	31108

```

; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 60/048,430
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1644)
; OTHER INFORMATION: Description of A
; -OTHER INFORMATION: Primer
US-09-086-663A-70

```

	Query Match	3.5%	Score 43.2;	DB 4;	Length 2294;
	Best Local Similarity	79.7%;	Pred.No. 0.028;		
	Matches 51;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;
QY	873	GCTGCTGCTGCTGCGCGTGTGTGTGTGCGCTGCAACTGCTGCTGCGGTGTGTGTCT	932		
Db	226	GCATGCTGCTGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	167		
QY	933	GCTG 936			
Db	166	GCTG 163			

```

RESULT 15
US-09-086-663A-1/c
; Sequence 1, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA
; APPLICANT: KARSENTY, GERARD
; TITLE OF INVENTION: OSG2/CBFA1 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UTSC:525
; CURRENT APPLICATION NUMBER: US/09/086,663A
; CURRENT FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/080,189
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 60/048,430
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-09-086-663A-1

```

	Query Match	3.5%;	Score 43.2;	DB 4;	Length 3334;	
	Best Local Similarity	79.7%;	Pred. No. 0.034;			
	Matches 51; Conservative	0;	Mismatches 13;	Indels 0;	Gaps 0;	
QY	873	GTGTCGCTGCTGCCCGCGTGGTTGTGGCTGCACACTGCTGCGCCTGTGTTGTTCT				932
Db	534	GTGTCGCTGCTGCTGCTGTTGTGCTGTTGCTGTTGCTGCTGCTGCTGCTGCT				475
QY	933	CGTG 936				
Db	474	CGTG 471				

Search completed: November 12, 2003, 23:23:39  
Job time : 84.0207 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 20:56:43 ; Search time 415.641 Seconds  
(without alignments)  
9519.735 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240

Sequence: 1 aggtgtgagtcagcccaaca.....gtaatacaactgtagtatag 1240

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270.6	21.8	474	11	US-09-918-995-3342
2	188.8	15.2	401	9	US-09-864-761-16305
3	140.4	11.3	398	10	US-09-983-965-4945
4	53.2	4.3	404	11	US-09-918-995-5026
5	50.4	4.1	609	12	US-10-027-632-43431
6	50.4	4.1	609	13	US-10-027-632-299775
7	50.4	4.1	609	13	US-10-027-632-43431
8	50.4	4.1	609	13	US-10-027-632-299775
9	50.2	4.0	650	12	US-10-027-632-201953
10	50.2	4.0	650	13	US-10-027-632-201953
11	49.6	4.0	606	12	US-10-027-632-134169
12	49.6	4.0	606	13	US-10-027-632-134169
13	47.6	3.8	77992	12	US-10-225-810-11
14	47.6	3.8	250000	12	US-10-225-810-26
15	47.6	3.8	1143	12	US-10-261-078-5
16	46.4	3.7	14152	9	US-09-764-869-1573

17	46.4	3.7	14152	14	US-10-091-504-1573	Sequence 1573, Ap
18	46.2	3.7	592	12	US-10-029-386-11128	Sequence 11128, A
19	46	3.7	2005	12	US-10-027-632-257598	Sequence 257598, A
20	46	3.7	2005	13	US-10-027-632-257598	Sequence 257598, A
21	46	3.7	3577	13	US-10-008-739A-1	Sequence 1, Appl
22	45.8	3.7	815	12	US-10-027-632-157282	Sequence 157282, A
23	45.8	3.7	815	13	US-10-027-632-157282	Sequence 157282, A
24	45.6	3.7	299	12	US-09-814-353-4844	Sequence 4844, Ap
25	45.6	3.7	299	13	US-09-814-353-4844	Sequence 1141, A
26	45.6	3.7	385	12	US-09-814-353-5368	Sequence 5368, Ap
27	45.6	3.7	385	13	US-09-814-353-5368	Sequence 11555, A
28	45.6	3.7	783	12	US-10-027-632-173913	Sequence 173913, A
29	45.6	3.7	783	13	US-10-027-632-173913	Sequence 173913, A
30	45.4	3.7	508	12	US-10-027-632-92930	Sequence 92930, A
31	45.4	3.7	508	13	US-10-027-632-92930	Sequence 92930, A
32	45.4	3.7	569	12	US-10-027-632-67529	Sequence 67529, A
33	45.4	3.7	569	13	US-10-027-632-295322	Sequence 295322, A
34	45.4	3.7	569	13	US-10-027-632-67529	Sequence 295322, A
35	45.4	3.7	637	12	US-10-027-632-316334	Sequence 316334, A
36	45.4	3.7	637	13	US-10-027-632-316334	Sequence 316334, A
37	45.4	3.7	637	13	US-10-027-632-316334	Sequence 316334, A
38	45.4	3.7	2145	12	US-10-027-632-86346	Sequence 86346, A
39	45.4	3.7	2145	13	US-10-027-632-86346	Sequence 86346, A
40	45	3.6	585	12	US-10-027-632-212466	Sequence 212466, A
41	45	3.6	585	13	US-10-027-632-212466	Sequence 212466, A
42	45	3.6	749	12	US-10-027-632-163055	Sequence 163055, A
43	45	3.6	749	13	US-10-027-632-163055	Sequence 163055, A
44	45	3.6	749	13	US-10-027-632-163055	Sequence 163055, A
45	45	3.6	749	13	US-10-027-632-163056	Sequence 163056, A

ALIGNMENTS

RESULT 1

Sequence 3342, Application US/09918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES  
FILE REFERENCE: 20411-756  
CURRENT APPLICATION NUMBER: US/09/918,995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235,076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3342  
LENGTH: 474  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(474)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-3342

Query Match	21.8%	Score	270.6	DB	11	Length	474
Best Local Similarity	98.6%	Pred. No.	1.5e-70				
Matches	273	Conservative	0	Mismatches	4	Indels	0
Gaps	0						
QY	964	ATTTCAAGAAATCTGAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAA	1023				
DB	54	ACTTTTAAAGAAATCTGAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAA	113				
QY	1024	TGAAATCTCCGGCTTACAAATTCAGATGAACAAAGACCAAGAACCGCTTCTCTCCCTCC	1083				
DB	114	TGAAATCTCCGGCTTACAAATTCAGATGAACAAAGACCAAGAACCGCTTCTCTCCCTCC	173				
QY	1084	CAATCTCTGTGATCTCAGTGTCTGAACAAAGAAAGTGTGTGGCCCTCTCTACCA	1143				

174 CAAATCTGTGAATCAGTATCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCACCA 233  
 1144 GGGGCTGATCAAGTGTCCAGGCGGAGGAGTCAATCCACAGGCTCTTTTAATCTGGC 1203  
 234 GGGGCTGATCAAGTGTCCAGGCGGAGGAGTCAATCCACAGGCTCTTTTAATCTGGC 293  
 1204 CAGTCTGAGAGTCAAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1240  
 294 CAGTCTGAGAGTCAAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 330

864-761-16305  
 ence 16305, Application US/09864761  
 at No. US2002048763A1  
 REAL INFORMATION:  
 PLICANT: Penn, Sharron G.  
 PLICANT: Rank, David R.  
 PLICANT: Hanzel, David K.  
 PLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 (LE REFERENCE: Aecomica-X-1  
 URGENT APPLICATION NUMBER: US/09/864,761  
 URGENT FILING DATE: 2001-05-23  
 RIOR APPLICATION NUMBER: US 60/180,312  
 RIOR FILING DATE: 2000-02-04  
 RIOR APPLICATION NUMBER: US 60/207,456  
 RIOR FILING DATE: 2000-05-26  
 RIOR APPLICATION NUMBER: US 09/632,366  
 RIOR FILING DATE: 2000-08-03  
 RIOR APPLICATION NUMBER: GB 24263.6  
 RIOR FILING DATE: 2000-10-04  
 RIOR APPLICATION NUMBER: US 60/236,359  
 RIOR FILING DATE: 2000-09-27  
 RIOR APPLICATION NUMBER: PCT/US01/00666  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00667  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00664  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00669  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00665  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00668  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00663  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00662  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00661  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: PCT/US01/00670  
 RIOR FILING DATE: 2001-01-30  
 RIOR APPLICATION NUMBER: US 60/234,687  
 RIOR FILING DATE: 2000-09-21  
 RIOR APPLICATION NUMBER: US 09/608,408  
 RIOR FILING DATE: 2000-06-30  
 RIOR APPLICATION NUMBER: US 09/774,203  
 RIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 16305  
 LENGTH: 401  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AF121782.1  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9  
 us-09-864-761-16305

Best Local Similarity 99.0%; Pred. No. 5.2e-46;  
 Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 115 AGGTTCTGGGCTCTGTAATCAAGTCAATAGAGGCCCCCAGAAATGCAACAGTCTCTGAAGG 174  
 Db 210 AGGTTCTGGGCTCTGTAATCAAGTCAATAGAGGCCCCCAGAAATGCAACAGTCTCTGAAGG 269  
 QY 175 CTCCAGGCTCTGTTCAACTGACGCTCTCCAGGCGTCCAGGCTGGAAGCTCATCATGTTGGCTCT 234  
 Db 270 CTCCAGGCTCTGTTCAACTGACGCTCTCCAGGCGTCCAGGCTGGAAGCTCATCATGTTGGCTCT 329  
 QY 235 CAGTGACATGCTGCTTAAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 294  
 Db 330 CAGTGACATGCTGCTTAAAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 389  
 QY 295 CACCTCTCAGAG 306  
 Db 390 CACCTCTCAGAG 401  
 RESULT 3  
 US-09-983-965-4945  
 ; Sequence 4945, Application US/09983965  
 ; Patent No. US20020137160A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Warren, Wesley C.  
 ; APPLICANT: Tao, Nengbing  
 ; APPLICANT: Byatt, John C.  
 ; APPLICANT: Mathialagan, Nagappan  
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 ; FILE REFERENCE: 37-21(10297)C  
 ; CURRENT APPLICATION NUMBER: US/09/983,965  
 ; CURRENT FILING DATE: 2001-10-26  
 ; PRIOR APPLICATION NUMBER: US 09/465,231  
 ; PRIOR FILING DATE: 1999-12-15  
 ; PRIOR APPLICATION NUMBER: US 60/113,678  
 ; PRIOR FILING DATE: 1998-12-17  
 ; NUMBER OF SEQ ID NOS: 5912  
 ; SEQ ID NO 4945  
 ; LENGTH: 398  
 ; TYPE: DNA  
 ; ORGANISM: Bos taurus  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 26-LIB34-017-Q1-E1-G9  
 US-09-983-965-4945  
 Query Match 11.3%; Score 140.4; DB 10; Length 398;  
 Best Local Similarity 78.5%; Pred. No. 1.9e-31;  
 Matches 169; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
 QY 113 GTAGTTCTGGGCTCTGTAATCAAGTCAATAGAGGCCCCCAGAAATGCAACAGTCTCTGAAG 172  
 Db 185 GCAGCTCTGTAATCAAGTCAATAGAGGCCCCCAGAAATGCAACAGTCTCTGAAG 244  
 QY 173 GGCTCCAGGCTCTGTTCAACTGACGCTCTCCAGGCGTCCAGGCTGGAAGCTCATCATGTTGGCT 232  
 Db 245 GGCTCCAGGCTCTGTTCAACTGACGCTCTCCAGGCGTCCAGGCTGGAAGCTCATCATGTTGGCT 304  
 QY 233 CTCAGTGACATGCTGCTTAAAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 292  
 Db 305 CTGAGAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 364  
 QY 293 TTCACCTCTCAGAGTACGACGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 326  
 Db 365 TTCACCTCTCAGAGTACGACGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 398  
 RESULT 4  
 US-09-918-995-5026  
 ; Sequence 5026, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: Hysed, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5026
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-5026

Query Match      4.1%; Score 53.2; DB 11; Length 404;
Best Local Similarity 68.9%; Pred. No. 3.1e-05;
Matches 73; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 TCAGTTTCCTAGCTGCCATAAAGCACCATAACCTGGTGGCTTAGAACAATGGAAG 87
Db 74 TGAGTTGCTGTGCTGCCATAAACAACACTATAGACTCTTTTGGCTTAACAATGGAAT 133

QY 88 GCATTTGCTCAGGTTCCAGAGCTGTAGGTTCTGGGTCGTGTAAT 133
Db 134 GTATTTCTCAGAGTTCCAGAGGCTGAAAGTCCAAAGATGCCACAAT 179

RESULT 5
US-10-027-632-43431
; Sequence 43431, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43431
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43431

Query Match      4.1%; Score 50.4; DB 12; Length 609;
Best Local Similarity 71.4%; Pred. No. 0.00028;
Matches 80; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 23 GTGATCAGTTTCG-TAGGCTGCATTAACAACACATAACTGTGGCTTAGAACAAT 81
Db 111 GTGATCAGTTTCCTTGGGCTGCCGTAACAAATTTACGACAACTTGTGGCTTAACAAT 170

QY 82 GGAAGGCATTTGCTCAGGTTCCAGAGCTGTAGGTTCTGGGTCGTGTAAT 133
Db 171 AGAAATGTATTCTCTCAAAGTTCTAGAGGCCAGAGTCTGAGATCCAGGCAT 222

RESULT 6
US-10-027-632-299775
; Sequence 299775, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 299775
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-299775

Query Match      4.1%; Score 50.4; DB 12; Length 609;
Best Local Similarity 71.4%; Pred. No. 0.00028;
Matches 80; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 23 GTGATCAGTTTCG-TAGGCTGCATTAACAACACATAACTGTGGCTTAGAACAAT 81
Db 111 GTGATCAGTTTCCTTGGGCTGCCGTAACAAATTTACGACAACTTGTGGCTTAACAAT 170

QY 82 GGAAGGCATTTGCTCAGGTTCCAGAGCTGTAGGTTCTGGGTCGTGTAAT 133
Db 171 AGAAATGTATTCTCTCAAAGTTCTAGAGGCCAGAGTCTGAGATCCAGGCAT 222

RESULT 7
US-10-027-632-43431
; Sequence 43431, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43431
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43431

```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43431
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43431

Query Match
Best Local Similarity 4.1%; Score 50.4; DB 13; Length 609;
Matches 80; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 23 GTGGATCAGTTCC-TAGGCTGCCATAACAAGACCACTAAGCTGGTGGCTTAGAACAAT 81
Db 111 GTGGATCAGTTCC-TAGGCTGCCATAACAAGACCACTAAGCTGGTGGCTTAGAACAAT 170

QY 82 GGAAGGCATTTCTCAGCGTTCCAGAGCTGTAGGTTCTGGTCTGGTAAAT 133
Db 171 AGAAATGTATTCTCTCAAGTTCTAGAGGCCAGAGTCTGAGATCCAGGCAT 222

RESULT 8
US-10-027-632-299775
; Sequence 299775, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 299775
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-299775

Query Match
Best Local Similarity 4.1%; Score 50.4; DB 13; Length 609;
Matches 80; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 23 GTGGATCAGTTCC-TAGGCTGCCATAACAAGACCACTAAGCTGGTGGCTTAGAACAAT 81
Db 111 GTGGATCAGTTCC-TAGGCTGCCATAACAAGACCACTAAGCTGGTGGCTTAGAACAAT 170

QY 82 GGAAGGCATTTCTCAGCGTTCCAGAGCTGTAGGTTCTGGTCTGGTAAAT 133
Db 171 AGAAATGTATTCTCTCAAGTTCTAGAGGCCAGAGTCTGAGATCCAGGCAT 222

RESULT 9
US-10-027-632-201953/c
; Sequence 201953, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201953
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201953

Query Match
Best Local Similarity 55.4%; Pred. No. 0.00033;
Matches 97; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 38 AGGCTGCCATAACAAGACCACTAAGCTGGTGGCTTAGAACAATGGAAGGCAATTTGCTC 97
Db 448 AGGCTGCCATAACAAGACCACTAAGCTGGTGGCTTAGAACAATGGAAGGCAATTTGCTC 389

QY 98 ACGTTCCAGAGCTGTAGGTTCTGGTCTGGTAAATCAAGTCAAGAGGCCCCAGAAAT 157
Db 388 ACAGTCCAGAGCTGTAGGTTCTGGTCTGGTAAATCAAGTCAAGAGGCCCCAGAAAT 329

QY 158 GCACAGTCTGAGGCTCCAGGCTCCAGGCTGGCTTCACTGACACCTCTCCAGGCT 212
Db 328 ACTCTATTCCTGGCTTCCAGACATTCATTTCTCACTGTGTGCTCACCAGCGTT 274

RESULT 10
US-10-027-632-201953/c
; Sequence 201953, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201953
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201953
```

Query Match 4.0%; Score 50.2; DB 13; Length 650;  
Best Local Similarity 55.4%; Pred. No. 0.00033;  
Matches 97; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 38 AGGCTGCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGGAAGGCATTGTCTC 97  
DB 448 AGGCTACCATACAAATGCCATAGACTGGGTGGCTTAAGACAGAAATTTATTTCTC 389

QY 98 ACGTTCAGAAAGCTGTAGTCTGGTCTGTGTAATGAAGTCATAGAAGGCCCCCAAGAT 157  
DB 388 ACAGTGCACAGAGCTGGAAGTCCAAAGATCAGGGTCCAGCATGGTTCAGGGGTCTGTGAAC 329

QY 158 GCAACAGTCTCCTAAGGGTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCT 212  
DB 328 ACTTATCTCTGGCTCCAGACATTCACCTTCTCACTGTGTCTCACCAGGCT 274

RESULT 11  
US-10-027-632-134169/c  
; Sequence 134169, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 134169  
; LENGTH: 606  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-134169

Query Match 4.0%; Score 49.6; DB 12; Length 606;  
Best Local Similarity 72.7%; Pred. No. 0.00048;  
Matches 64; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 24 TGGATCAGTTTCTAGGCTGCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGG 83  
DB 114 TGCATTAGTCCCTGGGGCTGCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGG 55

QY 84 AAGGCATTGTCTCAGGTTCCAGAGC 111  
DB 54 GAATTTATTCTCTCACCCTTCTGGAAC 27

RESULT 12  
US-10-027-632-134169/c  
; Sequence 134169, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 134169  
; LENGTH: 606  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-134169

Query Match 4.0%; Score 49.6; DB 13; Length 606;  
Best Local Similarity 72.7%; Pred. No. 0.00048;  
Matches 64; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 24 TGGATCAGTTTCTAGGCTGCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGG 83  
DB 114 TGCATTAGTCCCTGGGGCTGCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGG 55

QY 84 AAGGCATTGTCTCAGGTTCCAGAGC 111  
DB 54 GAATTTATTCTCTCACCCTTCTGGAAC 27

RESULT 13  
US-10-225-810-11/c  
; Sequence 11, Application US/10225810  
; Publication No. US20030157512A1  
; GENERAL INFORMATION:  
; APPLICANT: Berwinham, Jr., John R.  
; TITLE OF INVENTION: Transdorsins and Methods of Using Transdorsins  
; FILE REFERENCE: McLaugh-07165  
; CURRENT APPLICATION NUMBER: US/10/225,810  
; CURRENT FILING DATE: 2002-08-21  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11  
; LENGTH: 77992  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (25026)..(25026)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (34205)..(34304)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36809)..(36908)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (39455)..(39554)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (40803)..(40902)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:

```
; NAME/KEY: misc feature
; LOCATION: (42813)..(42912)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (43974)..(44073)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45116)..(45215)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (46542)..(46641)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47926)..(48025)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (49131)..(49230)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (70204)..(70303)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (76568)..(76667)
; OTHER INFORMATION: n is a, c, g, or t
; US-10-225-810-11
```

```
Query Match 3.8%; Score 47.6; DB 12; Length 77992;
Best Local Similarity 69.1%; Pred. No. 0.044;
Matches 65; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 28 TCAGTTCTCCTAGGCTCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGGAAG 87
Db 72074 TCATTTCCTGTTGCTGCTATACAAATACCAACAATTAGTAGATTAGAACAAATGCAAT 72015
QY 88 GCATTGCTCAGGTTCCAGAGCTGTAGTTCT 121
Db 72014 GTATTCTTATAGTTCTAGAGTCAGAGGTCT 71981
```

```
RESULT 14
US-10-225-810-26/c
; Sequence 26, Application US/10225810
; Publication No. US20030157512A1
; GENERAL INFORMATION:
; APPLICANT: Birmingham, Jr., John R.
; TITLE OF INVENTION: Transdorins and Methods of Using Transdorin
; FILE REFERENCE: McLaugh-07165
; CURRENT APPLICATION NUMBER: US/10/225,810
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 250000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (33774)..(33774)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (42953)..(43052)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45557)..(45656)
```

```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (48203)..(48302)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (49551)..(49650)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51561)..(51660)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (52722)..(52821)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (53864)..(53963)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55290)..(55389)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (56674)..(56773)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (57879)..(57978)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (78952)..(79051)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (85316)..(85415)
; OTHER INFORMATION: n is a, c, g, or t
; US-10-225-810-26
```

```
Query Match 3.8%; Score 47.6; DB 12; Length 250000;
Best Local Similarity 69.1%; Pred. No. 0.092;
Matches 65; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 28 TCAGTTCTCCTAGGCTCCATACAAAGCACCATACCTGGTGGCTTAGAACAAATGGAAG 87
Db 80822 TCATTTCCTGTTGCTGCTATACAAATACCAACAATTAGTAGATTAGAACAAATGCAAT 80763
QY 88 GCATTGCTCAGGTTCCAGAGCTGTAGTTCT 121
Db 80762 GTATTCTTATAGTTCTAGAGTCAGAGGTCT 80729
```

```
RESULT 15
US-10-261-078-5/c
; Sequence 5, Application US/10261078
; Publication No. US20030138954A1
; GENERAL INFORMATION:
; APPLICANT: TRONO, DIDIER
; APPLICANT: WIZNEROWICZ
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO LENTIVIRAL
; FILE REFERENCE: CLFR:01405
; CURRENT APPLICATION NUMBER: US/10/261,078
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1143
; TYPE: DNA
```

ORGANISM: Homo sapiens  
US-10-261-078-5

Query Match 3.8%; Score 47; DB 12; Length 1143;  
Best Local Similarity 68.4%; Pred. No. 0.0044;  
Matches 65; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 32 TTTCCTAGGCTGCCATACAAAGCACCATAACCTGGTGGCTTAGAACATGGAAAGGCAT 91  
DB 915 TTGACAGGGCTTCCATACAAAGTACCTCAGCCTGGATGTTTAAACACACAGAAATTTAT 856  
QY 92 TTGCTCAGGTTCCAGAGCTGTAGGTTCTGGGTC 126  
DB 855 TTTCACAGTTTGAAGCTACAAATTCAGATC 821

Search completed: November 13, 2003, 03:28:18  
Job time : 417.641 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: November 13, 2003, 03:28:26 ; Search time 15.3865 Seconds  
(without alignments)  
2412.580 Million cell updates/sec

Title: US-09-729-264-6  
Perfect score: 2077  
Sequence: 1 MERHLLTPVAVGSGSNEV.....HPQAFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76: \*  
1: pirl: \*  
2: pirl2: \*  
3: pirl3: \*  
4: pirl4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	151	7.3	404	1 I61596	advanced glycosyla
2	139.5	6.7	344	2 I56551	neurotrophin - rat
3	133	6.4	1427	2 I51669	tumor suppressor -
4	130	6.3	345	2 JC4025	opioid-binding cel
5	129.5	6.2	164	2 T24272	hypothetical prote
6	129	6.2	345	2 S03199	opioid-binding pro
7	129	6.2	6642	2 T29757	protein UNC-89 - C
8	128	6.2	345	2 JC1239	opioid-binding pro
9	128	6.2	1447	2 A54100	tumor suppressor p
10	127.5	6.1	871	1 I48696	protein-tyrosine k
11	127.5	6.1	881	1 I48697	protein-tyrosine k
12	125	5.9	1070	2 JC4593	protein-tyrosine k
13	122.5	5.9	188	2 T15851	hypothetical prote
14	122.5	5.9	4391	2 A38096	perlecan precursor
15	118.5	5.7	338	2 JC1238	opioid-binding pro
16	118.5	5.7	630	2 JH0593	Schwann cell myel
17	117.5	5.7	1443	2 I50600	neogenin - chicken
18	114.5	5.5	152	2 T18975	hypothetical prote
19	113.5	5.5	416	1 T18975	advanced glycosyla
20	111	5.3	693	2 S49228	sodium-dependent p
21	110.5	5.3	764	2 A49448	irregular chiasm C
22	110.5	5.3	1177	2 T16594	hypothetical prote
23	109.5	5.3	890	1 A53743	protein-tyrosine k
24	109	5.2	391	2 T09058	butyrophilin homol
25	108.5	5.2	487	2 S65133	butyrophilin - mou
26	106.5	5.1	802	2 T13149	mitogen-and stress
27	106.5	5.1	946	1 A47299	insulin-like recept
28	106.5	5.1	2491	1 A28372	neural cell adhesi
29	106	5.1	1092	1 JN0635	

differentiation an  
receptor tyrosine  
heparan sulfate pr  
hypothetical prote  
polyprotein - fava  
sdk protein - frui  
hypothetical prote  
connectin 3B - chi  
coxsackie- and ade  
sax-3 protein - Ca  
hypothetical prote  
CD22 homolog/B lym  
T-cell surface gly  
hypothetical prote  
killer cell inhibi  
50K glycoprotein p

30 105.5 5.1 862 2 I49583  
31 105.5 5.1 892 2 I38312  
32 105.5 5.1 3707 2 S18252  
33 105 5.1 662 2 T16525  
34 105 5.1 5825 2 T12117  
35 104 5.0 2222 2 T13924  
36 103.5 5.0 423 2 T29549  
37 103.5 5.0 1323 2 PNO568  
38 103 5.0 365 2 JC7780  
39 103 5.0 1273 2 T42405  
40 102 4.9 721 2 T41530  
41 102 4.9 868 2 A46512  
42 101.5 4.9 344 1 RWRTC2  
43 101.5 4.9 421 2 T46266  
44 101.5 4.9 841 2 JC5894  
45 101 4.9 338 2 JC5519

ALIGNMENTS

RESULT 1  
I61596  
advanced glycosylation end-products receptor precursor - human  
N/Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprote  
C/Species: Homo sapiens (man)  
C/Date: 24-May-1996 #sequence revision 07-Feb-1997 #text\_change 16-Jul-1999  
C/Accession: I61596; B42879; S27968  
R/Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko, H  
Genomics 23, 408-419, 1994  
A/Title: Three genes in the human MHC class III region near the junction with the clas  
nterpart of mouse mammary tumor gene int-3.  
A/Reference number: A55562; MUID:95137587; PMID:7835890  
A/Accession: I61596  
A/Status: nucleic acid sequence not shown; translation not shown; translated from GB/E  
A/Molecule type: DNA  
A/Residues: 1-404 <R>  
A/Cross-references: GB:D28769; NID:9561657; PIDN:BA05950.1; PID:9561659  
R/Neuper, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.;  
J. Biol. Chem. 267, 14998-15004, 1992  
A/Title: Cloning and expression of a cell surface receptor for advanced glycosylation  
A/Reference number: A42879; MUID:92340547; PMID:1378843  
A/Accession: B42879  
A/Molecule type: mRNA  
A/Residues: 'G', 2-99, 'R', 101-404 <NEE>  
A/Cross-references: EMBL:M91211; NID:G190845; PIDN:AAA03574.1; PID:G190846  
A/Experimental source: lung  
A/Note: sequence extracted from NCBI backbone (NCBIP:109438)  
C/Comment: Advanced glycosylation end products are heterogeneous nonenzymatically glyc  
cellular function, thus contributing to tissue lesions in diabetes.  
C/Comment: This receptor appears also to mediate the effects of amyloid beta peptide o  
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.  
C/Genetics:  
A/Gene: GDB:AGER  
A/Cross-references: GDB:306354; OMIM:600214  
A/Map position: 6p21.3-6p21.3  
A/Introns: 18/1; 53/3; 119/1; 140/3; 170/1; 231/1; 274/3; 322/1; 331/1; 373/2  
C/function:  
A/Description: neuronal receptor for amphoterin, a DNA-binding protein involved in neu  
C/Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
C/Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein  
F/1-22/Domain: signal sequence #status predicted <SIG>  
F/23-404/Product: advanced glycosylation end products receptor #status predicted <MAT>  
F/23-344/Domain: extracellular #status predicted <EXT>  
F/31-101/Domain: immunoglobulin homology <IM1>  
F/137-210/Domain: immunoglobulin homology <IM2>  
F/252-303/Domain: immunoglobulin homology <IM3>  
F/345-362/Domain: transmembrane #status predicted <TM>  
F/363-404/Domain: intracellular #status predicted <INT>  
F/25,81/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/38-99,144-208,259-301/Disulfide bonds: #status predicted

Query Match 7.3%; Score 151; DB 1; Length 404;

Best Local Similarity 23.3%; Pred. NO. 0.001;  
Matches 90; Conservative 38; Mismatches 112; Indels 146; Gaps 19;

QY 9 PEAVGSGSNEVIEGPQATVKGQARFNCFTVSG--WKLIWALSDMVLSVRPMEP 65  
Db 124 PIVLSAS--ELTAGVFN-----KVGTCVSEGSYPAGTSLSHLDG-----KP 163  
QY 66 IITNDRFTS-----ORYDQGNFT--SEMIHNVPSDSGNIR-----CSLQNSDLHGSAY 114  
Db 164 LPVNEKGVSKVQTRRHPTGLFTLQSELM---VTPARGSDPRPTFCSPGFLRHRAL 220  
QY 115 LTVQVWGELFIP---SVNLVVAENP-----CEVCLPSHWTRLPDLSWE 156  
Db 221 RTAPQPRVWEVPLEEVQLVV---EPGGAVAPGGTTLTCEVPAQS-----PQIHHM 272  
QY 157 LGLLVSHSHSYFVPSPDLSQAVSILATPQSGNLTCTVATWKSILKARKSATVNLTVIRC 216  
Db 273 KO-----GVPLFLPPSPVLLPEIGCPQDQGTYSVATHSSHGPOESRAVSISIE- 322  
QY 217 PDOTGGGINIPGVLSLPSGLSLPTWKGVLGLAGTMLLT-----PTCTLTIRCCC 269  
Db 323 PQEFG-----PTAGSVGSGGLTALALGILGLGTAAIIGVILW 363  
QY 270 RRRCCGNCNCCRCFCRRKRGFRIOFKKSEKVT--NKETETESGNENSGYNSDEOKT 327  
Db 364 QRR-----QRRGERKAPENQEBEERAEIN----- 389  
QY 328 TETASLPKSCSSDPEQRNNSCGPP 353  
Db 390 -----QSEEPAGESSTGGP 404

RESULT 2

I56551  
neurotrophin - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 19-May-2000  
C:Accession: I56551  
R:Struyk, A.F.; Canoll, P.D.; Wolfgang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.  
J. Neurosci. 15, 2141-2156, 1995  
A:Title: Cloning of neurotrophin defines a new subfamily of differentially expressed neur  
A:Reference number: I56551; MUID:95198094; PMID:7891157  
A:Accession: I56551  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-344 <RES>  
A:Cross-references: EMBL:U16845; NID:g755184; PIDN:AAA67445.1; PID:g755185  
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 6.7%; Score 139.5; DB 2; Length 344;  
Best Local Similarity 24.9%; Pred. NO. 0.0042;  
Matches 66; Conservative 40; Mismatches 110; Indels 49; Gaps 13;

QY 3 RHLTVPEAVGSGSNEVI-EGPQATVILKGSQARFNCFTVSGQWKLIMWALSDMVLS-- 59  
Db 20 RLLFLVPTGPVRSQDGFKAMDNVTVRQGESATILRCTIDNRVTRVAVLNRSTILYAGN 79  
QY 60 ----VRPMEPIITNDRFTSORYDQGNFTSEMIHNVPSDSGNIRCSLQ-----NSRL 109  
Db 80 DKWCLDFRVLLSN---TOTQY-----SRIQNVVDYDEGPTCSVQTDNHPKTSRV 128  
QY 110 HGSAYLTVQVWGELFIPFSLVNLVVAENPCEVCLPSHWTRLPDISWELGLLVSHSHSYFV 169  
Db 129 H----LIVQSPKPIVEISSIDISINEGNISLTCTATGRPE-FTVTWR---HISPKAVGFV 180  
QY 170 PPSDLSQAVSILATPQSGNLTCTVATWKSILKARKSATVNLTVIRCP-----QTTGGGI 224  
Db 181 SEDEYLE---IQGITREOSGEYECAS-NDVAAPVVRVAVVTVVAVPPYISEAKGTGVPV 235  
QY 225 NIPQVL-----SSLPSLGLFSLPTWGX 245  
Db 236 GQKGTIQCEASAVPSAEFQ----WFK 257

RESULT 3

I51669  
tumor suppressor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I51669  
R:Pierceall, W.B.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.  
Dev. Biol. 166, 654-665, 1994  
A:Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in t  
A:Reference number: I51668; MUID:95113183; PMID:7813784  
A:Accession: I51669  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1427 <PI>  
A:Cross-references: EMBL:U10986; NID:g606873; PIDN:AAA70168.1; PID:g606874  
C:Genetics:  
A:Gene: XDCCA

Query Match 6.4%; Score 133; DB 2; Length 1427;  
Best Local Similarity 25.7%; Pred. NO. 0.087;  
Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY 21 IEGPQATVILKGSQARFNCFTVSGQ--KLIMWALSDMVLSVRPMEPIITNDRFTSORYD 78  
Db 243 LQPSNVVAIEGQDAVLECAVS-GYPTPTIVWQGD-----EPVPIRTR---KYS 288  
QY 79 QGNGFTSEMIHNVPSDSGNIR--SLQNSRLHGSAYLTVQVWGELFIPFSLVNLVVAENE 136  
Db 289 VLGG--SNLLISNVTDGAGATCYCATYKNTSFSADLTVMVPPQFLNHPANLYAYESM 346  
QY 137 PCEVTCL----PSHWTRLPDISM-ELGLLVSHSHSYFVPEPSDLSQAVSILATPOSNT 191  
Db 347 DIEFECAVSGKPS-----PTVKWTKNGEVVPSDYFQIVDGSNLR---ILGLVKSDEGY 397  
QY 192 LTCVA-----TWKSLKARKSATVNLTVI-RCPODTGGGINIPGVLS 232  
Db 398 YQCIAENEAGNIQYLAQLIPDPVPSLSILPSAPRDV-----VPVLVSS 442

RESULT 4

JC4025  
opioid-binding cell adhesion protein - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 17-Mar-2000  
C:Accession: JC4025  
R:Shark, K.B.; Lee, N.M.  
Gene 155, 213-317, 1995  
A:Title: Cloning, sequencing and localization to chromosome 11 of a cDNA encoding a hu  
A:Reference number: JC4025; MUID:95237612; PMID:7721093  
A:Accession: JC4025  
A:Molecule type: mRNA  
A:Residues: 1-345 <SHA>  
A:Cross-references: GB:I34774; NID:g514373; PIDN:AAA36387.1; PID:g514374  
A:Experimental source: brain  
C:Comment: This protein binds opioid alkaloids in the presence of acidic lipids, exhib  
C:Genetics:  
A:Gene: GDB:OPCML; OBCAM; OPCM  
A:Cross-references: GDB:251677; OMIM:600632  
A:Map position: lppter-11qter  
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-term

Query Match 6.3%; Score 130; DB 2; Length 345;  
Best Local Similarity 25.6%; Pred. NO. 0.031;  
Matches 73; Conservative 34; Mismatches 108; Indels 70; Gaps 17;

QY 3 RHLTVPEAVGSGSNEVI-EGPQATVILKGSQARFNCFTVSGQWKLIMWALSDMVLS-- 58  
Db 20 RLLFLVPTGPVRSQDGFKAMDNVTVRQGESATILRCTIDNRVTRVAVLNRSTILYAG 78  
QY 59 ----SVRPMEPIITNDRFTSORYDQGNFTSEMIHNVPSDSGNIRCSLQ-----NSR 108  
Db 79 NDKWISIDPRVILVN---TPTQY-----SIMIQNVVDYDEGPTCSVQTDNHPKTSR 127

QY 109 LHGSAYLTVQVMGELFIPSVNLVVAENPEVETCL-----PSHWTRLPDISWELGLVSHS 164  
 Db 128 VH-----LIVQVPPQIMNISSDITVNEGSSVTLCLAIGRPE-----PVTWTR-----HL 172  
 QY 165 SYV-----FVPEPSDLOSASVILALTPQSNGLTLCVATWKSILKARKSATVNLTV-----IR 215  
 Db 173 SVKEQGQFVSEDEYLE-----ISDIKRDQSGEYECAL-NDVAAPDVRKVKITVNYPPYIS 227  
 QY 216 CPQDTGGGINIPGVL-----SSLPSLGFSLPTWKG-----VGLGLAG 252  
 Db 228 KAKNTGVSQKQILSCASAVPMAEFQ---WFKEDTRLATGLDG 269

## RESULT 5

T24272  
 Hypothetical protein T01B7.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T24272

R: Sims, M.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z19867

A:Accession: T24272

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-164 <WIL>

A:Cross-references: EMBL:Z66499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:T01B7.8

A:Experimental source: clone T01B7

C:Genetics:

A:Gene: CESP:T01B7.8

A:Map position: 2

A:Introns: 20/3; 90/2

Query Match 6.2%; Score 129.5; DB 2; Length 164;  
 Best Local Similarity 32.1%; Pred. No. 0.015;  
 Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;

QY 157 LGLLVSHSSYYFVPEPSDLOSASVILALTPQSNGLTLCVATWKSILKARKSATVNLTVIRC 216  
 Db 6 LAILLAIGTFIAV---SQVQSAV-----LPSVSTELATVGTDVSTASTAIDTLGNSSSRV 57

QY 217 PDPTGGGINIPGVLSLPSLGFSLPTWKGVLGLAGTMLTPTCTLTIRCCCRRCRCGC 276  
 Db 58 KQGGGCGCCGCGC-----GCGGCGGGGG---CGCCCRRCRCRCRCRCRCCTC 101

QY 277 --NCCC-RCFCRCR 287

Db 102 CRTCCCTRCCTCR 115

## RESULT 6

S03199

Oploid-binding protein OPCAM precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 17-Mar-2000

C:Accession: S03199

R: Schofield, P. R.; McFarland, K. C.; Hayflick, J. S.; Wilcox, J. N.; Cho, T. M.; Roy, S.; Le

EMO J. 8, 489-495, 1989

A:Title: Molecular characterization of a new immunoglobulin superfamily protein with po

A:Reference number: S03199; MUID:89251576; PMID:2721489

A:Accession: S03199

A:Molecule type: mRNA

A:Residues: 1-345 <SCH>

A:Cross-references: EMBL:X12672; NID:g585; PIDN:CAA31192.1; PID:g586

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

C:Keywords: transmembrane protein

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-345/Product: oploid-binding protein OPCAM #status predicted <MAT>

Query Match 6.2%; Score 129; DB 2; Length 345;  
 Best Local Similarity 25.7%; Pred. No. 0.037;  
 Matches 72; Conservative 33; Mismatches 109; Indels 66; Gaps 16;

QY 3 RHLLTYPEAVGSGSGNEVI-EGPQNAVTLKGSQARENCTVSGWKLIMWALSDMVVL----- 58  
 Db 20 RLLFLPTTGVVRSGDATPFKAMDNVTVRQGSATURCTIDRVTVRVAM-LNRSTILYAG 78  
 QY 59 ----SVRPMEPIITNDRFTLSQRYDOGNFTSEMIHNVPEPSDGNIRCSLQ-----NSR 108  
 Db 79 NDKWSIDPRVILLVN---TPTQY-----SIMIQNVVDVDEGPYTCVQTDNHPKTSR 127  
 QY 109 LHGSAYLTVQVMGELFIPSVNLVVAENPEVETCL-----PSHWTRLPDISWELGLVSHS 164  
 Db 128 VH-----LIVQVPPQIMNISSDITVNEGSSVTLCLAIGRPE-----PVTWTR-----HL 172  
 QY 165 SYV-----FVPEPSDLOSASVILALTPQSNGLTLCVATWKSILKARKSATVNLTV-----IR 215  
 Db 173 SVKEQGQFVSEDEYLE-----ISDIKRDQSGEYECAL-NDVAAPDVRKVKITVNYPPYIS 227  
 QY 216 CPQDTGGGINIPGVL-----SSLPSLGFSLPTWKGVLGLA 251  
 Db 228 KAKNTGVSQKQILSCASAVPMAEFQ---WFKEDTRLA 264

## RESULT 7

T29757

Protein UNC-89 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Dec-1999

C:Accession: T29757

R: Du, Z.; Le, T. T.; Wilson, R.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C09D1.

A:Reference number: Z20679

A:Accession: T29757

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-6642 <DUZ>

A:Cross-references: EMBL:AF003131; PIDN:AA854132.1; GSPDB:GN00019; CESP:unc-89

A:Experimental source: strain Bristol N2; clone C09D1

C:Genetics:

A:Gene: CESP:unc-89

A:Map position: 1

A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1;  
 3; 591/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match 6.2%; Score 129; DB 2; Length 6642;

Best Local Similarity 23.0%; Pred. No. 0.9;

Matches 59; Conservative 41; Mismatches 97; Indels 60; Gaps 10;

QY 6 LTVPEAVGSG-----SGNE-----VIEGPQNAVTLKGSQARENCT 40  
 Db 2038 LIIPNAQDSGKITVEASNEVSGSSSAQLTVNPPSTTPIVVDGPKSVTIKETETAEPKAT 2097

QY 41 VSQGW--KLIMWALSDMVLSVRPMEPIITNDRFTSQRYDOGNFTSEMIHNVPEPSDSG 98  
 Db 2098 IS-GFPAPTQKWTINEKIVEESRTITTKTEDVY-----LKISNAKIEQTG 2143

QY 99 NIRCLONSRLHGSAYLTVQVMGELFIPS-----VNLVVAENPEVETCL---PSHWTRL 150  
 Db 2144 TVKTAQNSAGQDSKQADLKVEPNKAPKFSQLTDKVADEGEPLRNWELDGSPCT-- 2201

QY 151 PDISWEL-GLLVSHSSYYFVPEPSDLOSASVILALTPQSNGLTLCVATWKSILKARKSATV 209  
 Db 2202 -EVSLLNGQLTKSDTVQVVDHGDGTGYHYTIAEAKPEMSGTLTAKAKNAAGCETSARKV 2260

QY 210 NLT-----VIRCPQD 219

Db 2261 TVNGNKKPFVQAPQN 2277

## RESULT 8

JC1239

Oploid-binding protein (clones SG8 and SG13) - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-May-2000

C/Accession: JCI1239

R/Lippman, D.A.; Lee, N.M.; Loh, H.H.

Gene 117, 249-254, 1992

A/Title: Opioid-binding cell adhesion molecule (OBAM) - related clones from a rat brain c

A/Reference number: JCI1238; MUID:92347701; PMID:11339369

A/Accession: JCI1239

A/Molecule type: mRNA

A/Residues: 1-345 <LIP>

A/Cross-references: GB:M88710; NID:G203247; PIDN:AAA40859.1; PID:G203248; GB:M88711; NID

C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

C/Keywords: transmembrane protein

Query Match 6.2%; Score 128; DB 2; Length 345;

Best Local Similarity 25.3%; Pred. No. 0.044;

Matches 73; Conservative 35; Mismatches 110; Indels 70; Gaps 17;

QY 3 RHLITVPEAVGSGGNEVI\_EGPNATVTLKGSQARFNCTVSGQKLMWALSDMVL--- 58

DB 20 RLLFLVPTGVPRSGDATFFPKMDNVTVRQGESATLRCITDDRVTRVAV-LNRSTILYAG 78

QY 59 ----SVRMEPIITNDRFTSORYDQGNFTSEMIHNVPSDSGNIRCSLQ-----NSR 108

DB 79 NDKWSIDPRVILVN---TPTQY-----SIMIQNVVDVDEGPTCSQTDNHPKTSR 127

QY 109 LHGSAYLTQVMGELFIPSNLVVAENPECEVTCL---PSHWTRLPDISWELGLVSHS 164

DB 128 VH----LIVQPPQIMNISDITVNEISSVTLCLAIQGRPE-----PTVTWR-----HL 172

QY 165 SYI-----VVPSPDIQSAVSLTALTPQNGTLTCVATWKSALKARFSATVNLTV-----IR 215

DB 173 SVKSGQGFVSDEYLE-----ISDTRQSGGYECSAL-NDVAAPDVVRKXITVNPYPYIS 227

QY 216 CPQDTGGGGINIPGVL-----SSLPLGFSPLTWGK-----VGLGLAGTML 255

DB 228 KAKTGVSVGQKGLSCASAVPMAEQ---WFKEDTSLATGLDGVRI 272

#### RESULT 9

A54100 tumor suppressor protein DCC precursor - human

N/Alternate names: colorectal cancer suppressor DCC

C/Species: Homo sapiens (man)

C/Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 05-Nov-1999

R/Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.

Genes Dev. 8, 1174-1183, 1994

A/Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.

A/Reference number: A54100; MUID:95011532; PMID:7926722

A/Accession: A54100

A/Molecule type: mRNA

A/Residues: 1-1447 <HED>

A/Cross-references: EMBL:X76132; NID:9453209; PIDN:CAA53735.1; PID:9453210

R/Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilto

Science 247, 49-56, 1990

A/Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.

A/Reference number: A40098; MUID:90100559; PMID:2294591

A/Accession: A40098

A/Molecule type: mRNA

A/Residues: 1-750 <FEA>

A/Cross-references: GB:M32292; NID:9181492; PIDN:AAA35751.1; PID:9181493

C/Genetics:

A/Gene: GDB:DCC

A/Cross-references: GDB:119838; OMIM:120470

A/Map position: 18q21.1-18q21.1

C/Keywords: transmembrane protein; tumor suppressor

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

F1-25/Domain: signal sequence #status predicted <SIG>

QY 14 SSGSNEV-----LEGPNATVTLKGSQARFNCTVSGW--KLIMWALSDM 55

DB 220 SKTGNEAEVRILSDPGLHRQLYFLQPSNVVAIEGKDAVECCVS-GYPPSFFTWLRGEE 278

QY 56 VVLSVRPMEPIITNDRFTSORYDQGNFTSEMIHNVPSDSGNIRCSLQNSRLHGS 113

DB 279 VI-----QLRSKKYSLGG--SNLLISNVTDDSGMYTCVTVTKENISASA 323

QY 114 YLTQVMGELFIPSNLVVAENPECEVTCLPSHWTRLPDISW-ELGLLVSSSYFYFPEP 172

DB 324 EUTLVPPFWFLNHPNSLNLYAYESMDIEFECTVS-GKPVPTVMKMGDVVIFSDIFQIVGG 382

QY 173 SLOQSAVSILALTPQNGTLTCVATWKSALKARFSATVNLTVIRCPQDTGGGINIPGVLS 232

DB 383 SNLR-----ILGVKSDGEFYQCVAENAGNAQTSQALIVPKPAIPSSS-----VLPS 430

QY 233 LP 234

DB 431 AP 432

#### RESULT 10

I48696

protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 2 - mouse

N/Alternate names: receptor-type tyrosine kinase

C/Species: Mus musculus (house mouse)

C/Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000

C/Accession: I48696; S60738

R/Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D.

Oncogene 11, 281-290, 1995

A/Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase

A/Reference number: I48696; MUID:95349951; PMID:7624144

A/Accession: I48696

A/Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-871 <GANI>

A/Cross-references: EMBL:X86444; NID:G929723; PIDN:CAA60165.1; PID:G929724

A/Experimental source: splice form 2

A/Accession: S60738

A/Molecule type: DNA

A/Residues: 1-456, 'A', 466-871 <GAN2>

A/Cross-references: EMBL:X86444; NID:G929723

A/Experimental source: splice form 4

C/Comment: For alternate splice forms see PIR:I48697.

C/Genetics:

A/Gene: nsk2

A/Cross-references: MGI:103308

C/Superfamily: mouse ror-related receptor; immunoglobulin homology; protein kinase hom

C/Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; tyr

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

F1-21/Domain: signal sequence #status predicted <SIG>

188	QY	SNGLTLCVAT-----WKSLKARKSATNLTIVR-----CPDDTGGGINTIPG-----V	222
278	DB	GLYTCIATNKHCKEFTAKAAATWSIAEWSKSQKDSQGYCAQYRGEGVLMOGPCKML	335
230	QY	LSSLPSPGFSPLP-----TWGKVGGLAGLTMLLTPTCIULTIRCCCRRCGCCCN--	277
336	DB	LVFLPTTSHRDPEDAQELLHTAWNEL-----KAVSPLCRPAEALLCYHLFLECSFG	388
278	QY	-----CCCCC-----FCCR	287
389	DB	VVETPMPICRBYCLAVKELFCAK	411

RESULT 12

protein-tyrosine kinase-related receptor PTK7 precursor - human  
N;Alternate names: receptor protein tyrosine kinase-like protein (RPTK)  
C;Species: Homo sapiens (man)  
C;Date: 16-Apr-1996 #sequence\_revision 24-May-1996 #text\_change 24-Sep-1999  
C;Accession: JG4593  
R;Park, S.K.; Lee, H.S.; Lee, S.T.  
J. Biochem. 119, 235-239, 1996  
A;Title: Characterization of the human full-length PTK7 cDNA encoding a receptor prote  
A;Reference number: JG4593; MUID:97037064; PMID:8862711  
A;Accession: JG4593  
A;Molecule type: mRNA  
A;Residues: 1-1070 <PAR>  
A;Cross-references: GR:U40271; NID:gl322231; PIDN:PAC50484.1; PID:gl322232  
C;Comment: This protein is a member of receptor protein tyrosine kinase family, but pr  
C;Genetics:  
A;Gene: GDB:PTK7  
A;Cross-references: GDB:134760; OMIM:601890

Query Match

Db 329 EPRVFTAGEE---RVTLCPKGLPEPSVWHEAGVRLPTEGRVY-----QKGHELVL 378

Qy 183 ALTPQSN-GTLTCVATWKS LKARKSATNLTV-----IRCPQDTGGGINTPGVLSL 233

Db 379 ANIAESDAGVYTHAA--NLAGRODYNIVATVPSWMLKKPPOSLEEGKPGVLDCL 434

RESULT 13

hypothetical protein C27A2.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T15651  
R:Nhan, M.  
submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid C27A2.

Db 322 IE-PBEG-----PTAGSVGGSLGTALALGILGLGTAALLIGV 361  
QY 267 CCRRCRCGNCRCRCFCRCRGRFIOQKKSEKKT--NKETETESGNSGNSDE 324  
Db 362 ILWQR-----QRRGEERKAPENQEBEERAEALN----- 390  
QY 325 QKTTDTASLPFKSCSSDPEQRNSSCGPP 353  
Db 391 -----QSEEPAGESSTGGP 405

## RESULT 2

US-10-184-644-559  
; Sequence 559, Application US/10184644  
; Publication No. US20030044930A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C227  
; CURRENT APPLICATION NUMBER: US/10/184,644  
; CURRENT FILING DATE: 2002-06-28  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 559  
; LENGTH: 2473  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-184-644-559

Query Match 6.6%; Score 138.5; DB 15; Length 2473;  
Best Local Similarity 29.5%; Pred. No. 0.015;  
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

QY 183 ALTPQSNGLTTCVATWKSLSKARKSATVNLTVIRCPQDTGGGINIPGVLSLSLGSFLT 242  
Db 2274 AATTGAAGTTTCAATTAATAATTTAATATGTTCC----- 2307  
QY 243 WGVGLGLAGTMLTPTCTLTIRCCCR-RRCCGNCRCRCFC 286  
Db 2308 -----ATTCTCATCGCCACCCACCCCGCCGCCACCC 2344

## RESULT 3

US-10-184-634-559  
; Sequence 559, Application US/10184634  
; Publication No. US20030068684A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C217  
; CURRENT APPLICATION NUMBER: US/10/184,634

; CURRENT FILING DATE: 2002-06-28  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 559  
; LENGTH: 2473  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-184-634-559

Query Match 6.6%; Score 138.5; DB 15; Length 2473;  
Best Local Similarity 29.5%; Pred. No. 0.015;  
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

QY 183 ALTPQSNGLTTCVATWKSLSKARKSATVNLTVIRCPQDTGGGINIPGVLSLSLGSFLT 242  
Db 2274 AATTGAAGTTTCAATTAATAATTTAATATGTTCC----- 2307  
QY 243 WGVGLGLAGTMLTPTCTLTIRCCCR-RRCCGNCRCRCFC 286  
Db 2308 -----ATTCTCATCGCCACCCACCCCGCCGCCACCC 2344

## RESULT 4

US-10-274-583-20  
; Sequence 20, Application US/10274583  
; Publication No. US20030138431A1  
; GENERAL INFORMATION:  
; APPLICANT: Exelixis, Inc.  
; TITLE OF INVENTION: LRRCAPS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-119C  
; CURRENT APPLICATION NUMBER: US/10/274,583  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: 60/361,196  
; PRIOR FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 1477  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-274-583-20

Query Match 6.6%; Score 138; DB 12; Length 1477;  
Best Local Similarity 25.1%; Pred. No. 0.0085;  
Matches 78; Conservative 38; Mismatches 127; Indels 68; Gaps 17;

QY 2 VAGAMENRDP-----GSGSGNEVIEGPNARVLKGSQARFNCTVS-OGWKLIWALS DMV 56  
Db 322 VAGEVKTQEVTLRYFGSPARPTFVIQPNTEVLGSEVTLECSATGHPPIRISWTRGDR 381  
QY 57 VLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNHVPDSGNIRCSLONS--RLHGSAY 114  
Db 382 PLVDPVNVITPS-----GG-----LYIQNVVQGSGEVACSATNIDSVHATAP 426  
QY 115 LTVQMGELFIPSVNLVVAENE-----PCEVTCTPSHWTMLPDISW-ELGLLVSHSSYFV 169  
Db 427 IIVQALPQFTVTPQDRVVIIEGQTVDFQCEAKGNPP-----PVIATWKGSLSDRAHLV 481  
QY 170 PEPSDQSAVSIALTPQSNGLTTCVATWKSLSKARKSATVNLTV-----IRCFQDT- 220  
Db 482 LSSGTLR--ISGVALHDQ--GOYEQAV--NIIGSKVVAHLTVQPRVTPVFASIESDTT 535  
QY 221 ---GGGINIPGVLSLSLGSFLP--TWGKVGGL--AGTMLLTPTTLTI----- 264  
Db 536 VEVGANVQLP-----CSSQGEPEPAITWKGQVGTBSGKFHISPEGFLTIINDVGPADAG 590  
QY 265 RCCCRRC 275  
Db 591 RYECVARNTIG 601





```

RESULT 11
US-10-180-410-8
; Sequence 8, Application US/10180410
; Publication No. US20030148382A1
; GENERAL INFORMATION:
; APPLICANT: SUN, CHAO
; APPLICANT: CARULLI, JOHN P.
; APPLICANT: LUKASHIN, ALEXANDER V.
; APPLICANT: KILBURN, DANIEL R.
; TITLE OF INVENTION: PANCREATIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A097 CIP
; CURRENT APPLICATION NUMBER: US/10/180,410
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: PCT/US01/19904
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/211,611
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-180-410-8

Query Match 6.1%; Score 128; DB 12; Length 708;
Best Local Similarity 23.1%; Pred. No. 0.023;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11

QY 14 SCSGNEVTEGPNARVLKGSQARFNCTVSQGWKLIMWLSVSRPMEPIITNDRT 73
DB 20 AGSPHFLOQPDIVVLGGEARLPCALGAYWGLVQWTKSLGALGGQR-----DLPG 71
QY 74 SORYQGGNFTS---EMIHNVPSDSGNTRCSLQNSRLHG-SAYITV-----QVME 122
DB 72 WSRYSWISGNAANGQDHLAIRPELEDEASYEQATAGLSRSRPAQLHLVLPPEAPQVLGG 131
QY 123 LFPISVNLVAENEPCEVTC-----LPSHWTWLPDISM-ELGLVSHSY--FVPE-- 171
DB 132 ---PSVSLVA---GVANLTCRSRGARPT-----PELLWFRDGVLLDGATHTLLKECT 181
QY 172 PSDLSQANSVILALTQSGNGLTCTVATWKSLSKARKSATVNLTVRCPODT 220
DB 182 PGSVESLTLPFSHDDGATFVCRARSOALPTGRDITLISLQYPPEVT 230

RESULT 12
US-10-199-672-584
; Sequence 584, Application US/10199672
; Publication No. US20030148442A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/199,672
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18

```

<p> ; PRIOR FILING DATE: 1997-10-17  ; PRIOR APPLICATION NUMBER: 60/063120  ; PRIOR FILING DATE: 1997-10-24  ; PRIOR APPLICATION NUMBER: 60/063121  ; PRIOR FILING DATE: 1997-10-24  ; PRIOR APPLICATION NUMBER: 60/063486  ; PRIOR FILING DATE: 1997-10-21  ; PRIOR APPLICATION NUMBER: 60/063540  ; PRIOR FILING DATE: 1997-10-28  ; PRIOR APPLICATION NUMBER: 60/063541  ; PRIOR FILING DATE: 1997-10-28  ; PRIOR APPLICATION NUMBER: 60/063542  ; PRIOR FILING DATE: 1997-10-28  ; PRIOR APPLICATION NUMBER: 60/063543  ; PRIOR FILING DATE: 1997-10-28  ; PRIOR APPLICATION NUMBER: 60/063544  ; PRIOR FILING DATE: 1997-10-28  ; Remaining Prior Application data removed - See File Wrapper or PALM.  ; NUMBER OF SEQ ID NOS: 612  ; SEQ ID NO 584  ; LENGTH: 708  ; TYPE: PRT  ; ORGANISM: Homo Sapien  US-10-187-749-584 </p>	<p> Query Match 6.1%; Score 128; DB 12; Length 708;  Best Local Similarity 23.1%; Pred. No. 0.023;  Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;  </p>
<p> QY 14 SSGSNEVIEGPNARVLKGSQARFNCTVSQGWKLIMWALSDMVLSVRPMEPIITNDRT 73  DB 20 AGSPHFLOQPEDLVLLGEEARLPCALGAYGLVQWTKSGALGGOR-----DLPG 71  QY 74 SRYDQGNFTS---EMITHNVPSDSGNIRCSLQNSRLHG-SAYLTV-----QWGE 122  DB 72 WRYWTSGNAAGQHDHLPVPELEDEASYEQATQAGLRSPALQLVLPPEAPQVLGG 131  QY 123 LFTPSNLVVAENPECEVTC-----LPSHWTWLPDISW-ELGLLVSHSSYV--FVPE-- 171  DB 132 ---PSVSLVA--GVPAULTCRSGDARPT-----PELLWFRDGVLLDGAFTHTLLKEGT 181  QY 172 PSDLSQSAVILALTPQSNGLTLCVATWKSILKARKSATVNLTVIRCPQDT 220  DB 182 PGSVESLTLPFFSHDDGATFVCRARSQALPTGRDTAITLSLQYPPEVT 230 </p>	<p> RESULT 14  US-10-194-457-584  ; Sequence 584, Application US/10194457  ; Publication No. US20030153037A1  ; GENERAL INFORMATION:  ; APPLICANT: Baker, Kevin P.  ; APPLICANT: Chen, Jian  ; APPLICANT: Desnoyers, Luc  ; APPLICANT: Goddard, Audrey  ; APPLICANT: Godowski, Paul J.  ; APPLICANT: Gurney, Austin L.  ; APPLICANT: Pan, James  ; APPLICANT: Smith, Victoria  ; APPLICANT: Watanabe, Colin K.  ; APPLICANT: Zhang, Zemin  ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  ; FILE REFERENCE: P3430R1C296  ; CURRENT FILING DATE: 2002-07-11  ; PRIOR APPLICATION NUMBER: 60/052586  ; PRIOR FILING DATE: 2002-01-15  ; PRIOR APPLICATION NUMBER: 60/059263  ; PRIOR FILING DATE: 1997-09-18  ; PRIOR APPLICATION NUMBER: 60/059266  ; PRIOR FILING DATE: 1997-09-18  ; PRIOR APPLICATION NUMBER: 60/062250  ; PRIOR FILING DATE: 1997-10-17  ; PRIOR APPLICATION NUMBER: 60/063120  ; PRIOR FILING DATE: 1997-10-24 </p>



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:40:51 ; Search time 4793.54 Seconds  
(without alignments)  
10582.571 Million cell updates/sec

Title: US-09-729-264-5  
Perfect score: 1240  
Sequence: 1 aggtgtgagtcagcaaca.....gtaatacaactgtagtatag 1240

Scoring table: IDENTITY NUC  
Gap 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1240	100.0	1240	6	AX380400 Sequence
2	1126.4	90.8	2051	9	AK092516 Homo sapi
3	1123.4	90.6	1175	6	AX380396 Sequence
4	1119	90.2	1168	6	AX380398 Sequence
5	1064	85.8	1139	6	AX380402 Sequence
6	380.6	30.7	1195	6	AX380404 Sequence
7	378.4	30.5	1155	10	AF537215 Mus muscu
8	318.6	25.7	142742	9	AF121782 Homo sapi
9	318.6	25.7	340000	9	AL163280 Homo sapi
10	187.4	15.1	149964	2	AC120145 Mus muscu
11	187.4	15.1	181510	2	AC120346 Mus muscu
12	187.4	15.1	196900	2	AC020851 Mus muscu
13	185.6	15.0	754	6	AX380408 Sequence
14	185.6	15.0	895	6	AX380406 Sequence
15	185.6	15.0	1556	10	BC004806 Mus muscu
16	153.6	12.4	47467	2	AC135456 Rattus no
17	153.6	12.4	140888	2	AC142238 Rattus no
18	149.8	12.1	181510	2	AC120346 Mus muscu
19	119	9.6	40205	9	AF045450 Homo sapi
20	119	9.6	170121	9	AF064860 Homo sapi
21	100	8.1	199665	9	AF064857 Homo sapi
22	58.4	4.7	143266	2	AC067790 Homo sapi
23	58.4	4.7	186968	9	AC007917 Homo sapi
24	58.2	4.7	149964	2	AC120145 Mus muscu
25	58.2	4.7	196900	2	AC020851 Mus muscu
26	56.6	4.6	427	9	HSU63332 Human super
27	56.2	4.5	163235	2	AC078808 Homo sapi
28	56.2	4.5	205933	9	AC117381 Homo sapi
29	55.6	4.5	250956	2	AC118358 Rattus no
30	55	4.4	1061	9	HSCAGCTG Homo sapi
31	55	4.4	123576	9	AC015969 Homo sapi
32	55	4.4	181616	2	AC144180 Macaca mu
33	54.2	4.4	138000	9	AC079127 Homo sapi
34	54.2	4.4	157743	2	AC067887 Homo sapi
35	54.2	4.4	183000	2	AC025134 Homo sapi
36	53.8	4.3	63692	9	AL603750 Human DNA
37	53.8	4.3	160541	2	AL359271 Homo sapi
38	53.8	4.3	206187	2	AC021172 Homo sapi
39	53.4	4.3	74057	9	AP005626 Homo sapi
40	53.4	4.3	164239	9	AC016074 Homo sapi
41	53	4.3	304785	2	AC118366 Rattus no
42	52.8	4.3	110000	3	AE003536 2 Continuation (3 of
43	52.8	4.3	138685	2	AC017582 Drosophil
44	52.8	4.3	166249	3	AC093502 Drosophil
45	52.8	4.3	185404	3	AC104703 Drosophil

ALIGNMENTS

RESULT 1  
AX380400  
LOCUS AX380400  
DEFINITION Sequence 5 from Patent WO0200710.  
ACCESSION AX380400  
VERSION AX380400.1 GI:19575330  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Welcher A.A., Sarmiento J.M., Schultz, H.J. and Chute, H.T.  
TITLE B7-like molecules and uses thereof  
JOURNAL Patent: WO 0200710-A 5 03-JAN-2002;

linear PAT 18-MAR-2002

Amgen, Inc. (US)	
Location/Qualifiers	
1. .1240	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
80. -1240	
/note="unnamed protein product"	
/codon_start=1	
/protein_id="CAD28218.1"	
/translation="MERLLTVPEAVGSGSNEVIEGPQNAVILKGSQARPNCTVSGS	
WKLIMWLSMDVMSVSRWMEPIITNDRTSQYDGGNFTSEMIHNVPSDSNIR	
SLONRSLHSAVLTQVMGELFIPSNVLVAENPEVTCIPSHWTRLPDLSWELGL	
VSHSVYFVPEPDLQSAVSLATPQSNGLTLCVATWKSILKARKSAVNLVTRCPQ	
DYGGINIPGVLSLPSLPSLPTAGKVLGLAGTMLPTCTLTIRCCCRRCRCQ	
NCCRCFCPCRRKRGFRIFQFKSEKEKINKEKETESNENSGNSDEOKTETASIP	
PKSCSSDPEORNSGCPHQRADQRRPRPASHPQAFNLASPEKVSNTVV"	
BASE COUNT	318 a 319 c 305 g 298 t
ORIGIN	
Query Match	100.0%; Score 1240; DB 6; Length 1240;
Best Local Similarity	100.0%; Pred. No. 4.4e-283;
Matches 1240; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AGGTGTGAGTCCAGCAACAGTGTGGATCAGTTTCTAGCTGCGCATACAAAGCACCAT 60
Db	1 AGGTGTGAGTCCAGCAACAGTGTGGATCAGTTTCTAGCTGCGCATACAAAGCACCAT 60
Qy	61 AACCTGGTGGCTTAGAACCAATGAAAGGCAATTTGCTCAGGTTCCAGAAAGCTGTAGGTTTC 120
Db	61 AACCTGGTGGCTTAGAACCAATGAAAGGCAATTTGCTCAGGTTCCAGAAAGCTGTAGGTTTC 120
Qy	121 TGCGTCTGGTAAATGAAGTCAATGAAGGCCCCCAGAAATGCAACAGTCTGAAAGGGTCCCA 180
Db	121 TGCGTCTGGTAAATGAAGTCAATGAAGGCCCCCAGAAATGCAACAGTCTGAAAGGGTCCCA 180
Qy	181 GGTCTGCTTCACTGACACGCTCTCCAGAGGCTGGAGCTCATCATGTGGGCTCTCAGTGA 240
Db	181 GGTCTGCTTCACTGACACGCTCTCCAGAGGCTGGAGCTCATCATGTGGGCTCTCAGTGA 240
Qy	241 CATGTGGTGGTCTAAGCGTTCAGGCCCCATGAGGCCCCATCATCAACCAATGACCGCTTCACCTC 300
Db	241 CATGTGGTGGTCTAAGCGTTCAGGCCCCATGAGGCCCCATCATCAACCAATGACCGCTTCACCTC 300
Qy	301 TCAGAGGTACGACAGGCGGAGAACTTCACCTCGGAGATGATCATCCACAATGTGGAGCC 360
Db	301 TCAGAGGTACGACAGGCGGAGAACTTCACCTCGGAGATGATCATCCACAATGTGGAGCC 360
Qy	361 CAGTGAATTCGGGGAACATCAGATGACAGCTCCAGAACAGTCCGCTGCGATGATCTGCTTA 420
Db	361 CAGTGAATTCGGGGAACATCAGATGACAGCTCCAGAACAGTCCGCTGCGATGATCTGCTTA 420
Qy	421 CCTTACGCTCCAGTTATGGAGAGCTGTTCATCCAGTGTAACTTGTAGTGGCTGA 480
Db	421 CCTTACGCTCCAGTTATGGAGAGCTGTTCATCCAGTGTAACTTGTAGTGGCTGA 480
Qy	481 GAATGAACCTTGTGAAGTACTTGTCTACCTCCACTGGAACCGGCTCCCGGATATTTC 540
Db	481 GAATGAACCTTGTGAAGTACTTGTCTACCTCCACTGGAACCGGCTCCCGGATATTTC 540
Qy	541 CTGGAGCTCGGTCTCTGTGAGCACTTCAAGCTATTATTTGTTCCGAGCCGAGGA 600
Db	541 CTGGAGCTCGGTCTCTGTGAGCACTTCAAGCTATTATTTGTTCCGAGCCGAGGA 600
Qy	601 CCTTCAAGTGCAGTGAACATCTCGGCTCTGACCCACAGAGCAATGGGACTTTGACTTG 660
Db	601 CCTTCAAGTGCAGTGAACATCTCGGCTCTGACCCACAGAGCAATGGGACTTTGACTTG 660
Qy	661 CGTGGCTACCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAACTGTAAATCTCACTGTGAT 720
Db	661 CGTGGCTACCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAACTGTAAATCTCACTGTGAT 720
Qy	721 TCGGTCTCCCAAGACACATGGAGGTGATTAATATATTCAGGTGTATTATCAAGTTTACC 780
Db	721 TCGGTCTCCCAAGACACATGGAGGTGATTAATATATTCAGGTGTATTATCAAGTTTACC 780
Qy	781 GAGTTTAGGTTTTTCAATTCCTACTTGGGGCAAAAGTTGGACTTGGACTAGCAGGACCAT 840
Db	781 GAGTTTAGGTTTTTCAATTCCTACTTGGGGCAAAAGTTGGACTTGGACTAGCAGGACCAT 840
Qy	841 GCTTCTGACGCGACGCTGACTCTTACAAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db	841 GCTTCTGACGCGACGCTGACTCTTACAAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy	901 TGCTCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db	901 TGCTCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy	961 TCAATTTCAAAGAAATCTGAAAAGAGAGCAACAAAGAACTGACACAGAAAGTGG 1020
Db	961 TCAATTTCAAAGAAATCTGAAAAGAGAGCAACAAAGAACTGACACAGAAAGTGG 1020
Qy	1021 AAATGAAACTCCGGCTACAAATTCAGATGAAACAAAGACACAGAAACCGCTTCTCTCCC 1080
Db	1021 AAATGAAACTCCGGCTACAAATTCAGATGAAACAAAGACACAGAAACCGCTTCTCTCCC 1080
Qy	1081 TCCCAAAATCCTGTGAATCCAGTATCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCA 1140
Db	1081 TCCCAAAATCCTGTGAATCCAGTATCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCA 1140
Qy	1141 CCAGCGGGCTGATCAACAGTCCACCCAGGCGCAGCAAGTCAATCCACAGGCTTCTTTAATCT 1200
Db	1141 CCAGCGGGCTGATCAACAGTCCACCCAGGCGCAGCAAGTCAATCCACAGGCTTCTTTAATCT 1200
Qy	1201 GGCAGCTCTGAGAGTCACTAGTAAATCACTGTAGTATAG 1240
Db	1201 GGCAGCTCTGAGAGTCACTAGTAAATCACTGTAGTATAG 1240
RESULT 2	
AK092516	2051 bp mRNA linear PRI 15-JUL-2002
LOCUS	Homo sapiens cdna FLJ35197 fis, clone PLACE6017788, highly similar to IGSF5.
DEFINITION	AK092516
ACCESSION	AK092516.1 GI:21751130
VERSION	AK092516
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 Katakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
AUTHORS	NEDO human cDNA sequencing project
TITLE	Unpublished
JOURNAL	2 (bases 1 to 2051)
REFERENCE	Isogai, T. and Yamamoto, J.
AUTHORS	Direct Submission
TITLE	Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
JOURNAL	Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES	Location/Qualifiers	Db	1032	AAAGAGGATTCGTTATTCATTTTCAAAAGAAATCTGAAAAGAGAGACAAACGAAA	109
source	1..2051	QY	1005	CTGAGACAGAAAGTGGAAATGAAACCTCCGGCTCAATTCAGATGAAACAAAGACACAG	1064
	/organism="Homo sapiens"	Db	1092	CTGAGACAGAAAGTGGAAATGAAACCTCCGGCTCAATTCAGATGAAACAAAGACACAG	1151
	/mol_type="mRNA"	QY	1065	AAACCGCTCTCTCCCTCCCAATCCCTGTGAATCCAGTGATCCCTGAAACAAGAAACAGTA	1124
	/db_xref="taxon:9606"	Db	1152	AAACCGCTCTCTCCCTCCCAATCCCTGTGAATCCAGTGATCCCTGAAACAAGAAACAGTA	1211
	/clone="PLACE6017788"	QY	1125	GCTGTGGCCCTCTCTCACCGCGGCTGATCAACGTCCACCGCCGACGCAAGTCATCCAC	1184
	/tissue_type="placenta"	Db	1212	GCTGTGGCCCTCTCTCACCGCGGCTGATCAACGTCCACCGCCGACGCAAGTCATCCAC	1271
	/clone_lib="PLACE6"	QY	1185	AGCTTCTTTTAAATCTGGCCAGTCTCTGAGAAGGTCTGTAATACAACTGTAGTATAG	1240
	/note="cloning vector: pWE18SFL3"	Db	1272	AGCTTCTTTTAAATCTGGCCAGTCTCTGAGAAGGTCTGTAATACAACTGTAGTATAG	1327
BASE COUNT	580 a 463 c 474 g 534 t				
ORIGIN					
Query Match	90.8%; Score 1126.4; DB 9; Length 2051;				
Best Local Similarity	99.5%; Pred. No. 3.8e-256;				
Matches 1130; Conservative	0; Mismatches 6; Indels 0; Gaps 0;				
QY	105 CAGAAGCTGAGGTTCTGGGTCTGGTAATGAAGTCATGAAGGCCCCAGAAATGCAACAG	164			
Db	192 CAGTGTGGACGTTCTGGGTCTGGTAATGAAGTCATGAAGGCCCCAGAAATGCAACAG	251			
QY	165 TCTGTAGGGCTCCACGGCTCGCTTCAATGCACCGTCTCCACGGCTGGAGGTCATCA	224			
Db	252 TCTGTAGGGCTCCACGGCTCGCTTCAATGCACCGTCTCCACGGCTGGAGGTCATCA	311			
QY	225 TGTGGGCTCTCAGTGACATGGTGTCTAAGCGTCAGGCCCATGAGCCCATCATCACCA	284			
Db	312 TGTGGGCTCTCAGTGACATGGTGTCTAAGCGTCAGGCCCATGAGCCCATCATCACCA	371			
QY	285 ATGACCGCTTCACTCTCAGAGTACACACAGGGGGGAATTCACCTGGAGATGATCA	344			
Db	372 ATGACCGCTTCACTCTCAGAGTACACACAGGGGGGAATTCACCTGGAGATGATCA	431			
QY	345 TCACAAATGTGAGCCAGTGATTCGGGGAACATCAGATGCAGCCTCCAGAACAGTCGCC	404			
Db	432 TCACAAATGTGAGCCAGTGATTCGGGGAACATCAGATGCAGCCTCCAGAACAGTCGCC	491			
QY	405 TGAATGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGGTGTTCATTCAGTGTTA	464			
Db	492 TGAATGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGGTGTTCATTCAGTGTTA	551			
QY	465 ATCTTGTAGTCGTGAGATGAACCTTGTGAAGTACTTGTCTACCTCCACATGAGGCC	524			
Db	552 ATCTTGTAGTCGTGAGATGAACCTTGTGAAGTACTTGTCTACCTCCACATGAGGCC	611			
QY	525 GGCTCCCGGATAATTCTCGGAGCTCGGTCTCTCGTCTAGCCATTCAGCTATTTTG	584			
Db	612 GGCTCCCGGATAATTCTCGGAGCTCGGTCTCTCGTCTAGCCATTCAGCTATTTTG	671			
QY	585 TTCCGAGGCCAGCGACCTTCAAAGTGCAGTGCAGTCTGCTGCTGACCCACACAGCA	644			
Db	672 TTCCGAGGCCAGCGACCTTCAAAGTGCAGTGCAGTCTGCTGCTGACCCACACAGCA	731			
QY	645 ATGGGACTTTGACTTGGCTGGCTACCTTGAAGAGCCTGAAGGCCCGCAAGTCTGCAATG	704			
Db	732 ATGGGACTTTGACTTGGCTGGCTACCTTGAAGAGCCTGAAGGCCCGCAAGTCTGCAATG	791			
QY	705 TAATCTCAGTGATTCGGTGTCCCAAGACACTGGAGTGGTATTAATTCAGGTG	764			
Db	792 TAATCTCAGTGATTCGGTGTCCCAAGACACTGGAGTGGTATTAATTCAGGTG	851			
QY	765 TAATATCAAGTTTACCAGTTTATAGTTTTTTCATGTCTACTTGGGGCAAGTTGACTTTG	824			
Db	852 TAATATCAAGTTTACCAGTTTATAGTTTTTTCATGTCTACTTGGGGCAAGTTGACTTTG	911			
QY	825 GACTAGAGGACCATGCTTCTGACGGCGAGTGTTACTCTTACATATACGCTCTGCTGT	884			
Db	912 GACTAGAGGACCATGCTTCTGACGGCGAGTGTTACTCTTACATATACGCTCTGCTGT	971			
QY	885 GCCGCGCTGTGTGTGGTGTCAAATGCTGCTGCGGTGTGTGTGTCTCTGCTGAGAA	944			
Db	972 GCCGCGCTGTGTGTGGTGTCAAATGCTGCTGCGGTGTGTGTGTCTCTGCTGAGAA	1031			
QY	945 AAAGAGGATTTCTGTATTTCAAAAGAAATCTTGAAAAGAGAGAGACAAACAAAGAA	1004			

[illegible]

RESULT 4				
AX380398	AX380398	Sequence 3 from Patent WO0200710.	1168 bp	DNA
LOCUS				
DEFINITION				
ACCESSION	AX380398			
VERSION	AX380398.1	GI:19575328		
KEYWORDS	.			
SOURCE	Homio sapiens (human)			
ORGANISM	Homio sapiens			
				linear
				PAT 18-MAR-2002

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Patent: WO 0200710-A 3 03-JAN-2002;  
 Angen, Inc. (US)

FEATURES  
 source  
 1. .1168  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 8. .1168  
 /notes="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD28217.1"  
 /db\_xref="GI:19575329"  
 /translation="MVVAGMENRDPGSGSNGNEVIEGPONARVLKGSQARFNCTVSGOG  
 WKLIWALSDMVVLSVRPMEITINDRFTSQYDQGNFTSEMIHNHVPDSGNTSRC  
 SLQNSLRGSAYLTQVQMGVSLTIPSLVAENEPCEVTLPSHMTWLPDLSWEGLGL  
 VSHSSYFVPEPSDILQASVILATLPSQNGTLTCVATKSLKARKATVNLTVIRCPQ  
 DTGGNIGVLLGSLGVSFSLTQKSGVLGAGTMLLPTCTLTITRCCCRRCRCGCC  
 NCCRCRCRRKRGRIQFKKSEKNTKMETETESGNENSGYNSEDEQKTTDTASLP  
 PKSCSSDEQRNSSCGPHQADRPFRPASHPQAFSLNAPKSVNTTVV"

BASE COUNT 299 a 302 c 288 g 279 t

ORIGIN

Query Match	30.2%;	Score 1119;	DB 6;	Length 1168;
Best Local Similarity	99.1%;	Pred. No. 2.1e-254;		
Matches 1125;	Conservative	0;	Mismatches 10;	Indels 0; Gaps 0;
QY 106	AGAAGCTGTAGGTTCTGGGTTCTGGTAAATGAAGTCATAGAAGGCGCCCCAGAAATGCAACAGT	165		
DB 34	AGACCCACCGGTTCTGGGTTCTGGTAAATGAAGTCATAGAAGGCGCCCCAAATGCAAGAGT	93		
QY 166	CTTGAAGGGCTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGGCTGGGAAGCTCATCAT	225		
DB 94	CCTGAAGGGCTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGGCTGGGAAGCTCATCAT	153		
QY 226	GTGGGCTCTCAGTGACATGTGTGTCTAAGCGTCAGGCCCATGAGCCCATCATCACAA	285		
DB 154	GTGGGCTCTCAGTGACATGTGTGTCTAAGCGTCAGGCCCATGAGCCCATCATCACAA	213		
QY 286	TGACCGGCTTCACTCTCAGAGGTACACACAGGGCGGGAACTTCACCTCGGAGATGATCAT	345		
DB 214	TGACCGGCTTCACTCTCAGAGGTACGACACAGGGCGGGAACTTCACCTCGGAGATGATCAT	273		
QY 346	CCAAATGTGGAGCCAGTGATTCGGGGAACATCATGACAGCTCCAGAACAGTCCGCT	405		
DB 274	CCAAATGTGGAGCCAGTGATTCGGGGAACATCATGACAGCTCCAGAACAGTCCGCT	333		
QY 406	GCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATCCCAAGTGTAA	465		
DB 334	GCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATCCCAAGTGTAA	393		
QY 466	TCCTTGAGTCTGAGAAATGAAACCTTGTAAGTTACTTGTCTACCTTCACATGGAACCG	525		
DB 394	TCCTTGAGTCTGAGAAATGAAACCTTGTAAGTTACTTGTCTACCTTCACATGGAACCG	453		
QY 526	GCTCCCGGATATTCCTGGGAGCTCGGTCTCTCTGTCAGCCATTCAGCTATTATTTTGT	585		
DB 454	GCTCCCGGATATTCCTGGGAGCTCGGTCTCTCTGTCAGCCATTCAGCTATTATTTTGT	513		
QY 586	TCCGGAGCCCGAGGACCTTCAAAAGTCAGTGAGCATCTCGGCTCTGACCCACAGAGCAA	645		
DB 514	TCCGGAGCCCGAGGACCTTCAAAAGTCAGTGAGCATCTCGGCTCTGACCCACAGAGCAA	573		
QY 646	TGGACTTTGACTTGCTGCTACCTGGAGAGGCTGGAAGCCCGGAAGTCTGCAACTGT	705		
DB 574	TGGGACTTTGACTTGCTGCTGCTACCTGGAGAGGCTGGAAGCCCGGAAGTCTGCAACTGT	633		
QY 706	AAATCTCAGTGTGATTCGGTGTCCCAAGACACTGGGAGGTGGTATTAAATATTCAGGTGT	765		

Db 634 AAATCTCACTGTGATTCGGTCTCCCAAGACACTGGAGGTGTATTAATATTCAGGTGT 693

Qy 766 ATTATCAAGTTTACGAGTTTAGTGTTCCTATGCTTACTTGGGCAAGATTGACATTGG 825

Db 694 ATTATCAAGTTTACGAGTTTAGTGTTCCTATGCTTACTTGGGCAAGATTGACATTGG 753

Qy 826 ACTAGCAGGCAACCATGCTTCTGACGCCGACGCTGCTTCTTACAATACGCTGCTGCTG 885

Db 754 ACTAGCAGGCAACCATGCTTCTGACGCCGACGCTGCTTCTTACAATACGCTGCTGCTG 813

Qy 886 CCGCGTCTGTTGTTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945

Db 814 CCGCGTCTGTTGTTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 873

Qy 946 AAGAGGATTCGTTATCAATTTCAAAAGAAATCTGAAAAGAGAGAACAAACAAAGAAC 1005

Db 874 AAGAGGATTCGTTATCAATTTCAAAAGAAATCTGAAAAGAGAGAACAAACAAAGAAC 933

Qy 1006 TGAGACAGAAAGTGAATGAAATCGGCTCAAAATTCAGATGAAACAAAGACCAAGAAC 1065

Db 934 TGAGACAGAAAGTGAATGAAATCGGCTCAAAATTCAGATGAAACAAAGACCAAGAAC 993

Qy 1066 AACCGTCTCTCTCCCTCCCAATCTGTAATCCAGTATCCTGAAACAAAGACCAAGAAC 1125

Db 994 CACCGCTTCTCTCCCTCCCAATCTGTAATCCAGTATCCTGAAACAAAGACCAAGAAC 1053

Qy 1126 CTGTGGCCCTCTCCACAGCGGGTGTATCAACGCTCCACCGGCGAGCAAGTCAATCCACA 1185

Db 1054 CTGTGGCCCTCTCCACAGCGGGTGTATCAACGCTCCACCGGCGAGCAAGTCAATCCACA 1113

Qy 1186 GCCTTCTTTTATCTGGCCAGTCTCTGAGAGAGTCAAGTCAATCAATCAATCAATCAAT 1240

Db 1114 GCCTTCTTTTATCTGGCCAGTCTCTGAGAGAGTCAAGTCAATCAATCAATCAATCAAT 1168

RESULT 5

AX380402

LOCUS AX380402 1139 bp DNA linear PAT 18-MAR-2002

DEFINITION Sequence 7 from Patent WO0200710.

ACCESSION AX380402

VERSION AX380402.1 GI:19575332

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

WEICHER, A., SARMIENTO, U.M., SCHULTZ, H.J. and CHUTE, H.T.

AUTORS B7-like molecules and uses thereof

TITLE Patent: WO 0200710-A 7 03-JAN-2002;

JOURNAL Amgen, Inc. (US)

FEATURES

Location/Qualifiers

1..1139

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

1..1134

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD28219.1"

/db\_xref="GI:19575332"

/translation="MVAGAMENRDPGSGSNEVIEGPNARVLKGSQARFNCTVSQG

WKLIMWLSDMVLSVRPMELITNDRTSQRVDGGLNLTSEMIHNVEPSDSNIR

SLONSLGSAVLTQVNGELFIPSNLVVAENPEVTCPLPSHWRILPDISWELGL

VSHSYVFPSPDLQSAVSLALTQSNGLTCTVATWKSILKARKSTVNLTVIRCPQ

DTGGINIPGVISLPSLGSFSLPTWGRVGLGAGMLTPTCTITFCCRRRCGCG

NCCGCCFCKRRLNKKRKQKLEKEMKTPATIQMKRPQTLLSLPNPVPV

LINKETVAALLTGLNINHPGQVHRLLLIWFVLRSLVIQL"

Matches 1102; Conservative 0; Mismatches 10; Indels 22; Gaps 1;

Qy 106 AAGAGCTGTAGGTTCTGGTCTGGTAAATGAAGTCATAGAGGCCCGCAGAAATGCAAGT 165

Db 27 AGACCCACCCGGTCTGGGTCTGGTAAATGAAGTCATAGAGGCCCGCAGAAATGCAAGT 86

Qy 166 CTTGAAGGGCTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCAT 225

Db 87 CTTGAAGGGCTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCAT 146

Qy 226 GTGGGCTCTCAGTACATGCTGGTGTAGCTAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 285

Db 147 GTGGGCTCTCAGTACATGCTGGTGTAGCTAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 206

Qy 286 TGACCGCTTCACTCTCAGAGGTAGACACAGGCGGGAATCTTCACTCGGAGATGATCAT 345

Db 207 TGACCGCTTCACTCTCAGAGGTAGACACAGGCGGGAATCTTCACTCGGAGATGATCAT 266

Qy 346 CCAATATGTGGAGCCAGTATTTCGGGGAAATCAGATGCGCTTCCAGAAACAGTGGCT 405

Db 267 CCAATATGTGGAGCCAGTATTTCGGGGAAATCAGATGCGCTTCCAGAAACAGTGGCT 326

Qy 406 GCATGAGTCTGCTTACCTTACCGTCCAGTATTGGAGAGCTGTTCAATCCAGTGTAA 465

Db 327 GCATGAGTCTGCTTACCTTACCGTCCAGTATTGGAGAGCTGTTCAATCCAGTGTAA 386

Qy 466 TCTTGTAGTCTGCTGAGAAATCTTGTGAAGTACTTGTCTACCTCACAACCTGAGCCG 525

Db 387 TCTTGTAGTCTGCTGAGAAATCTTGTGAAGTACTTGTCTACCTCACAACCTGAGCCG 446

Qy 526 GCTCCCGGATATTTCTGGAGCTCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585

Db 447 GCTCCCGGATATTTCTGGAGCTCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506

Qy 586 TCCGGAGCCCGAGCCTTCAAAGTGCAGTGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 645

Db 507 TCCGGAGCCCGAGCCTTCAAAGTGCAGTGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 566

Qy 646 TGGAGCTTTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705

Db 567 TGGAGCTTTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626

Qy 706 AAATCTCACTGTGATTCGGTCTCCCAAGACACTGGAGTGTATTAATATTCAGGTGT 765

Db 627 AAATCTCACTGTGATTCGGTCTCCCAAGACACTGGAGTGTATTAATATTCAGGTGT 686

Qy 766 ATTATCAAGTTTACCGAGTTTAGGTTTTTCAATTCGCTTCTGGGGCAAAAGTTGGA 825

Db 687 ATTATCAAGTTTACCGAGTTTAGGTTTTTCAATTCGCTTCTGGGGCAAAAGTTGGA 746

Qy 826 ACTAGCAGGCAACCATGCTTCTGACGCCGAGCTGTACTCTTACATACGCTGCTGCTGCTG 885

Db 747 ACTAGCAGGCAACCATGCTTCTGACGCCGAGCTGTACTCTTACATACGCTGCTGCTGCTG 806

Qy 886 CCGCCGTCTGTTGTGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945

Db 807 CCGCCGTCTGTTGTGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866

Qy 946 AAGAGGATTCGTTATCAATTTCAAAGAAATCTGAAAAGAGAGAACAAACAAAGAAC 1005

Db 867 AAGAG-----GAAATCTGAAAAGAGAGAACAAACAAAGAAC 904

Qy 1006 TGAGACAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1065

Db 905 TGAGACAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 964

Qy 1066 AACCGTCTCTCCCTCCCAATCTCTGTAATCCAGTGTATCTGAAACAAAGAAACAGTAG 1125

Db 965 CACCGCTTCTCTCCCTCCCAATCTCTGTAATCCAGTGTATCTGAAACAAAGAAACAGTAG 1024

Qy 1126 CTGTGGCCCTCTCCACAGGCGGCTGATCAAGTCTCCAGCCAGGCGAGGAGTCAATCCACA 1185

Db 1025 CTGTGGCCCTCTCCACAGGCGGCTGATCAAGTCTCCAGCCAGGCGAGGAGTCAATCCACA 1084

BASE COUNT

290 a 300 c 283 g 266 t

ORIGIN

Query Match

Best Local Similarity 85.8%; Score 1064; DB 6; Length 1139;

97.2%; Pred. No. 2.3e-241;

QY 1186 GGCTCTTTTAACTGCGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATA 1239  
Db 1085 GGCTCTTTTAACTGCGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATA 1138

RESULT 6  
AX380404 AX380404 1195 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 9 from Patent WO0200710.  
ACCESSION AX380404  
VERSION AX380404.1 GI:19575334  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus

REFERENCE 1  
AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.  
TITLE B7-like molecules and uses thereof  
JOURNAL Patent: WO 0200710-A 9 03-JAN-2002;  
Amgen, Inc. (US)

FEATURES  
source Location/Qualifiers  
1..1195  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
53..1165  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD28220.1"  
/db\_xref="GI:19575335"  
/translation="MEGSHRDVLAIVLILQALFASGSSYQIIEGPQNVVLKDESHF  
NCTVTHGKLMWTLQWVLSLTQGPITNNRFTYASYNDSFISELIHIDVPS  
DSGVQCSLONSHGFGAFLSQVMGTINIPNNLIIVTEGEPQNVTCYAVGWSLFDI  
SWELEVPVSHSSYNSFLPEGNFVMSVLDLTPLNGTGLTCAELKLOAQKSLTVNL  
TVQPPDSIGEEGALPTWAILLAVAFSLILILVLIIFCCCAASREKEESTY  
QNEYRKSNWRNTKADPETKLKSKEYNGYSSDEAKAAQTASLPPKSAEVSLEPKRSS  
SLPYGLNKHOPGATHPRVSRFDSAFKVRNVLV"

CDS  
313 a 304 c 291 g 287 t

BASE COUNT 313 a 304 c 291 g 287 t

ORIGIN  
Query Match 30.7%; Score 380.6; DB 6; Length 1195;  
Best Local Similarity 62.4%; Pred No. 2.3e-79;  
Matches 702; Conservative 0; Mismatches 354; Indels 59; Gaps 4;

QY 115 AGTTTCTGGGTCGTGAATGAAGTCATAGAGGCCCCCAGAAATGCAACAGTCCTGAAGGG 174  
Db 109 AGCTTCCGGATCCAGTTATCAGATCATAGAGGTCCTCAGAAATGTAACAGTCCTTAAAGGA 168

QY 175 CTCACAGGCTCGCTTCAACTGACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCT 234  
Db 169 CTCAGAGCTCACTTCAACTGACCGTGAATCATCGGCTGGAGCTTCTCATGTGGACTCT 228

QY 235 CAGTGACATGTGTGTCGTAAGCGTCAGGCCCATGAGCCCATCATCATCAACATGACCGCTT 294  
Db 229 TAACCAATGTGTGTGTCGTAAGCTCTCACACCAAGGACCCATCATCATCAACACCGCTT 288

QY 295 CACCTCTCAGAGTACGACCGGCGGAACCTTCACTCGGAGATGATCATCCACAATGT 354  
Db 289 CACCTATGCCAGTTAAACAGACTGACAGCTTATCTCGGAGTTGATCATCCATGATGT 348

QY 355 GGAGCCCATGTGATTCGGGGAACATCAGATGACGACCTCCAGAACATGTCCTGCAATGGATC 414  
Db 349 GCAGCCCATGTGATTCGGGATCCGTGCAATGACGCTCGAGAACATGTCCTGCAATGGATC 408

QY 415 TGCCTACCTTACCTCCAGTTATCGGAGAGCTGTTCAATCCAGTGTATATCTGTAGT 474  
Db 409 TGCCCTTCCCTCAGTGCAGAGTCATGGGACCCCTGAACATCTCTAGCAACACCTTATAGT 468

QY 475 CGCTGAGATGAACCTTGTGAGTGTACTTGTCTACCTCTCACACTGGAGCCCGGCTCCCGGA 534  
Db 469 CACTGAGGCTGAACCTGTGATGTGACTTGTGCTATGCTATGCTGCGGTGGCTGCACTCACTCCCGGA 528

QY 535 TATTTCTCGGAGCTCGCTCTCTGCTCAGCCATTCAAGCTATTATTTTGTTCGGAGCC 594  
Db 529 TATTTCTCGGAGCTCGCTCTCGGAGCTTCCGTTAGCCATTCCGAGTTACAACTCTTCTGGAGCC 588

QY 595 CAGCGACCTTCAAAAGTCAGTGAGCATCTCGCTCTGACCCCAAGAGCAATGGAGCTTT 654  
Db 589 GGGCAACTTTATGAGGGTCTTGGAGTCTCTGACCTCACACCACTGGGCAACGGGACCTT 648

QY 655 GACTTCCGCTGCTACTCTGGAAGAGCTGCAAGCCGCGAAGTCTGCACTGTAAATCTCAC 714  
Db 649 GACTTGTGGCAGAGCTGAAGGACTTTCAGCCCAAGAGTCTTAACTGTCAACCTGAC 708

QY 715 TGTGATTCGGTCTCCCAAGACACTGGAGGTGGTATTAAATATTCCAGGTGTATTATCAAG 774  
Db 709 TGTGCTTCCAGCTCTCCACTGACAGTATTGGAGAGGA----- 744

QY 775 TTTACCGAGTTAGGTTTTTTCATTCGCTTACTTGGGGCAAGTTGGACTTGGACTAGCAGG 834  
Db 745 -----AGGCCAGCAGCTGCCGACCTGGGCCATCATCTGCTGGCAGTGGCCTT 792

QY 835 CACCATGCTTCTGACGCGCAGCTGTAATCTTCAATACGCTGCTGCTGCTGCGCCGCTCG 894  
Db 793 TTCTTGTCTTGTATCTCTGATCATTTGTTTGTATTATATATTTCTGTTGCTG----- 843

QY 895 TTGTTGTGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954  
Db 844 ----TGTGCTCTCCAGGAGAGAGAAAGAGAAATCTACTTATCAAAATGAAATAGGAAT 899

QY 955 TCGTATTCAATTTCAAAAGAAATCTGAAAGAGAGAGACAAACAAAGAAACTTGAGACAGA 1014  
Db 900 CTGCAAAACATGAGGACAA-----ACAAAGCAGATCCGGAGACAAAGTTAAA 945

QY 1015 AAGTGAATGAAACTCGGCTCAAACTCAGATGAAACAAAGACCAAGAAACCGCTTC 1074  
Db 946 AAGTGAAGAGGAAACTACGGTACAGTTCCGATGAGGCAAGGCTGCACACTGCACTC 1005

QY 1075 TCTCTCTCCCAATCTGTGATCCAGTATCTCTCAACAAAGAAAGAGTGTGCTGCTGCGCC 1134  
Db 1006 TCTCTCTCTCTAAATCTGTGAAGTCAAGCTTCCAGAAACCGACAGTA-----GCCT 1059

QY 1135 TCTCACCAGCGGCTGATCAACGCTCCACCCAGGCGCAGCACTCATCCAGGCTTCTTT 1194  
Db 1060 TCTTATCAGGAACATCAATAAATCAGCCCGTCCAGCACTCATCCAGGCTTCTCTT 1119

QY 1195 TAATCGGCACTCTGAGAGGTCAGTAATACAACTGTAGTATA 1239  
Db 1120 TGACATCGCCAGTCTCAGAGGTCAGAAATGTGACTTTAGTGA 1164

RESULT 7  
AF537215 AF537215 1155 bp mRNA linear ROD 03-JUN-2003  
LOCUS Mus musculus cell adhesion molecule JCAM mRNA, complete cds.  
DEFINITION AF537215 GI:31339733  
ACCESSION AF537215.1  
VERSION AF537215.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus

REFERENCE 1  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Hirabayashi, S., Tajima, M., Yao, I., Nishimura, W., Mori, H. and Hata, Y.  
TITLE JAM4, a Junctional Cell Adhesion Molecule Interacting with a Tight Junction Protein, MAGI-1  
JOURNAL Mol. Cell. Biol. 23 (12), 4267-4282 (2003)  
MEDLINE 22658521  
PUBMED 12773569  
REFERENCE 2 (bases 1 to 1155)  
AUTHORS Hata, Y., Hirabayashi, S. and Tajima, M.  
TITLE Direct Submission  
JOURNAL Submitted (10-AUG-2002) Medical Biochemistry, Tokyo Medical and

Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan  
 Location/Qualifiers  
 1. .1155  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Swiss Webster"  
 /db\_xref="taxon:10090"  
 43. .1155  
 /note="interacts with MAGI-1; localized at tight junctions in epithelial cells"  
 /codon\_start=1  
 /product="cell adhesion molecule JCAM"  
 /protein\_id="AAP49218.1"  
 /db\_xref="GI:31339734"  
 /translation="MEGSWRDVLAVILVLAQLTASGSSYQIIEGPQWVTVLKDESAHF  
 NCTVHGWKLWTLNQWVLSLTTQGPFIITNNRFTYASINSDSFISELILIDVQPS  
 DSGVQCSLQSHGFSAPLVSQVMTLNIPNNLI VTEGPEQNVTCYAVGWSLSDPI  
 SWELPVPDSISGSPALPTWAILLAVAFSILLIILVLIIFCCCASTRKEESTY  
 QNEIRKSAMRTNKADPTKLGKKNYGSDEAKAAQTASLPKSAEVSLEPKRSS  
 SLPYDELAKHQGPATPRVDFDIASPKVENVTLV"  
 299 a 296 c 283 g 277 t

BASE COUNT 299 a 296 c 283 g 277 t

ORIGIN  
 Query Match 30.5%; Score 378.4; DB 10; Length 1155;  
 Best Local Similarity 62.3%; Pred. No. 7.6e-79;  
 Matches 701; Conservative 0; Mismatches 356; Indels 69; Gaps 4;

QY 115 AGGTTCTGGTCTGTAATAGTATCATAGAGGCCCCCAGAGTCAACAGATCCCTGAAGGG 174  
 DB 99 AGCTCCGGATCCAGTTATCATCATAGAGGTCTCAGATGTAAACGTCTTAAAGGA 158  
 QY 175 CTCACAGGCTCGCTTCAACTGACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGCTCT 234  
 DB 159 CTCAGAGGCTCACTTCAACTGACCGTGTACTCAAGCTGGAAGCTTCTCATGTGGACTCT 218  
 QY 235 CAGTGACATGGTGTCTAAGCTGAGGCCCATGAGCCCATCATCAACCAATGACCGCTT 294  
 DB 219 TAACCAAAATGGTGTGCTGAGTCTACACCCCAAGGACCATCATCAACCAACCGCTT 278  
 QY 295 CACCTCTCAGAGGTACGACAGCGGGGAACTTCACTCCGAGATGATCATCCACAATGT 354  
 DB 279 CACCTATGCCAGTTACAACAGACACTGACAGCTTCATCTCGGAGTTGATCATCCATGATG 338  
 QY 355 GAGGCCAGTATTCGGGAACTCAGATGAGCTCCAGACACTCCGCTGCTGATGATC 414  
 DB 339 GCAGCCCAAGTACTCGGATCCGTCGAATGCAAGCTTCGAGACACCCATGGTGGATC 398  
 QY 415 TGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATTCCTCCAGTGTAACTTGTAGT 474  
 DB 399 TGCTTTCCTCTCAGTCCAAAGTATGAGGACCTGGAACATTCCTAGCAACACCTTATAGT 458  
 QY 475 CGCTGAGATGACCTTGTGAGTACTTGTCTACCTCACACTCGACCGGCTCCCGGA 534  
 DB 459 CACTGAGGGGAAACCTGTAAATGTGACTGTCTATGCGGTGGGCTGAGCTCACTCCCGGA 518  
 QY 535 TATTTCTGGGAGCTCGGCTCTCTGCTGAGCACTTCAAGCTATATTTTGTTCGGAGCC 594  
 DB 519 TATTTCTGGGAGCTTGGGTTCCCGTAAAGCAATTCGAGTTACAATTCCTTCTGGAGTC 578  
 QY 595 CAGGACCTTCAAGTGCAGTGAGCATCTGGCTCTGACCCACACAGCAATGGGACTTT 654  
 DB 579 GGGCAACTTATGAGGGTCTTGTAGTGTCTGTGACCTTCAACACCTGGGCAAGGGACCTT 638  
 QY 655 GACTTGGTGGCTACCTTGAAGAGCTGGAAGGCCCGCAAGTCTGCAACTGTAAATCTCAC 714  
 DB 639 GACTTGTGTGGCAGAGCTGAAGGACTTGCAGGCCACCAAGTCTTAACTGTCAACTGAC 698  
 QY 715 TGTGATTCGGTGTCCCAAGACACTGAGGTGTGTATTAATATTCAGGTTGATTAAG 774  
 DB 699 TGTGGTTTCAGCTCCCACTGACAGTATGAGAGGA----- 734  
 QY 775 TTACCAGATTAGGTTTTTTTCAATGCTACTTGGGGCAAGTTGGACTTGGACTAGCAGG 834

Db 735 -----AGCCCCAGCAGCTGCCGACCTGGCCATCATCTGCTGGCAGTGGCCTT 782  
 QY 835 CACCATGTTCTGAGCGCAGCGTGTACTCTTTACAAATACGCTGCTGCTGCCCGCTCG 894  
 Db 783 TTCTTGTCTCTTGATCTCTGATCATTTGTTTGAATATAATTTCTGTTGCTG----- 833  
 QY 895 TTGTTGTGGCTGCAACTGCTGCTGCGTGTGTTTCTGCTGTAGAGAGAGAAAGAGGATT 954  
 Db 834 -----TTGTGCTCTCCAGGAGAGAAAGAGAGAAATCTACTTATCAAAATGAATAGGAAT 889  
 QY 955 TCGTATTCAATTTCAAAAGAAATCTGAAAAAGAGAGACAAACAAAGAAATCTGAGACGA 1014  
 Db 890 CTGCAAAACATGAGGACAA-----ACAAAGCAGATCCGGAGAGACAAAGTTAAA 935  
 QY 1015 AAGTGGAAATGAAACTCCGCTCAATTCAGATGAACAAAGAACCCACAGAAACCGCTTC 1074  
 Db 936 AGTGGAAAGGAAACTACGGGTACGTTCCGATGAGCAAGGCTGACAGATGCTATC 995  
 QY 1075 TCTCCCTCCCAATCTGTGAATCCAGTGATCCCTGAAACAAAGAAACAGTAGCTGTGGCC 1134  
 Db 996 TCTCCCTCTCTAAATCTGCTGAAGTCAGCTTCCAGAAAAAGCCAGCAGTA-----GCCT 1049  
 QY 1135 TCTCACCAGCGGCTGATCAAGTCCACCCAGGCCAGCAGTCCATCCACAGGCTTCTTT 1194  
 Db 1050 TCTTATCAGGAATCAATATAACATCAGCCGGTCCAGCAACTCATCCCGGTTTCTT 1109  
 QY 1195 TAATCTGGCCAGTCTCTGAGAAGTCAAGTAAATCACTAGTATAGTAG 1240  
 Db 1110 TGACATCGCCAGTCTCAGAGGTCAAGAAATGTGACTTTAGTAGTAG 1155

RESULT 8  
 AF121782  
 LOCUS Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.  
 DEFINITION AF121782  
 ACCESSION AF121782  
 VERSION AF121782.1 GI:4210991  
 KEYWORDS HIG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 142742)  
 AUTHORS Taudien,S., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,  
 Schattervoy,R., Weber,J., Schilling,M., Menzel,U., Yaspo,M.L. and  
 Rosenthal,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular  
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

FEATURES  
 Location/Qualifiers  
 1. .142742  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.3"  
 /clone="PAC 206A10"  
 3. .118  
 /evidence=not experimental  
 /rpt\_family="MIR"  
 complement(140..415)  
 /evidence=not experimental  
 /rpt\_family="MLT2E"  
 complement(668..881)  
 /evidence=not experimental  
 /rpt\_family="MSR58A"  
 complement(1078..1162)  
 /note="GRAIL, score = 86.000%, comment = excellent shadow"  
 /evidence=not experimental  
 1114..1403  
 /evidence=not experimental  
 /rpt\_family="MLT1I"

repeat\_region  
 3. .118  
 /evidence=not experimental  
 /rpt\_family="MIR"  
 complement(140..415)  
 /evidence=not experimental  
 /rpt\_family="MLT2E"  
 complement(668..881)  
 /evidence=not experimental  
 /rpt\_family="MSR58A"  
 complement(1078..1162)  
 /note="GRAIL, score = 86.000%, comment = excellent shadow"  
 /evidence=not experimental  
 1114..1403  
 /evidence=not experimental  
 /rpt\_family="MLT1I"

repeat_region	1481..1603 /evidence=not_experimental /rpt_family="L2"	repeat_region	9966..10041 /note="homology = 100.00%, score = 38, counts = 2"
exon	1576..1710 /note="MZEF, score = 66.1%" /evidence=not_experimental	repeat_region	10244..10664 /evidence=not_experimental /rpt_family="MLTIE"
exon	1576..1710 /note="GRAIL, score = 99.000%, comment = excellent"	repeat_region	10665..11028 /evidence=not_experimental /rpt_family="MLTIF"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIF"	repeat_region	11318..11417 /evidence=not_experimental /rpt_family="L1MC/D"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIF"	repeat_region	11503..11818 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /note="GRAIL, score = 50.000%, comment = good shadow"	exon	11503..11818 /evidence=not_experimental /rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIF"	repeat_region	12902..13180 /evidence=not_experimental /rpt_family="MLTIC"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIF"	repeat_region	12902..13180 /evidence=not_experimental /rpt_family="MLTIC"
exon	1576..1710 /note="GRAIL, score = 65.000%, comment = good shadow"	exon	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /evidence=not_experimental /rpt_family="HERVL"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /note="MZEF, score = 95.1%" /evidence=not_experimental	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /note="GenScan, score = 1.83%, comment = initial_exon 6 bp frame: 1 phase: 0"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIC"	exon	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIC"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIC"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental /rpt_family="MLTIC"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /note="GRAIL, score = 45.000%, comment = marginal"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /note="GRAIL, score = 71.000%, comment = good"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"
exon	1576..1710 /note="GRAIL, score = 81.000%, comment = excellent shadow"	repeat_region	13180..13180 /evidence=not_experimental /rpt_family="AluSg"

```

/evidence=not_experimental
complement(18397..18584)
/feature="GRAIL, score = 66.000%, comment = good"
/evidence=not_experimental
complement(18643..24647)
/evidence=not_experimental
/rpt_family="HERV17"
complement(19228..19361)
/feature="WZEF, score = 92.8%"
/evidence=not_experimental
20900..21135
/feature="WZEF, score = 51.4%"
/evidence=not_experimental
21150..21275
/feature="Xpound exon prediction, score = 74% (0%)"

Query Match      25.7%; Score 318.6; DB 9; Length 142742;
Best Local Similarity 98.8%; Pred. NO. 1.5e-64;
Matches 321; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 115 AGTCTTGGTCTGTAAGTATGATGATAGTATGATAGAGGCCCCAGAGTCAACAGTCCCTGAAGGG 174
DB 44273 AGTCTTGGTCTGTAAGTATGATGATAGAGGCCCCAGAGTCAACAGTCCCTGAAGGG 44332

QY 175 CTCACAGGCTCGCTTCAACTGACACCGCTCTCCAGGGCTGGAAGTCAATCATGTCGGCTCT 234
DB 44333 CTCACAGGCTCGCTTCAACTGACACCGCTCTCCAGGGCTGGAAGTCAATCATGTCGGCTCT 44392

QY 235 CAGTGACATGNGGTTGCTAAGCGTCAGGCCCATGAGCCCATCATCAACATGACCCGCTT 294
DB 44393 CAGTGACATGNGGTTGCTAAGCGTCAGGCCCATGAGCCCATCATCAACATGACCCGCTT 44452

QY 295 CACCTCTCAGAGTACGACACCGGGCGGAACCTTCACTCGAGATGATCAATCAATGTC 354
DB 44453 CACCTCTCAGAGTACGACACCGGGCGGAACCTTCACTCGAGATGATCAATCAATGTC 44512

QY 355 GAGGCCAGTATGTCGGGACATCAGATGAGCGCTCCAGACAGTCCGCTGCGATGATC 414
DB 44513 GAGGCCAGTATGTCGGGACATCAGATGAGCGCTCCAGACAGTCCGCTGCGATGATC 44572

QY 415 TGCTTACCTTACCGTCCAGTTATG 439
DB 44573 TGCTTACCTTACCGTCCAGTTATG 44597

RESULT 9
HS21C080      340000 bp      DNA      linear      PRI 24-MAY-2000
LOCUS      Homo sapiens chromosome 21 segment HS21C080.
DEFINITION      AL163280 AP001735 BA000005
ACCESSION      AL163280.2 GI:7717369
VERSION
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 340000)
Hattori.M., Fujiyama.A., Taylor.T.D., Watanabe.H., Yada.T.,
Park.H.S., Toyoda.A., Ishii.K., Totoki.Y., Choi.D.K., Soeda.E.,
Ohki.M., Takagi.T., Sakaki.Y., Taudien.S., Blechschmidt.K.,
Polley.A., Menzel.U., Delabar.J., Kumpf.K., Lehmann.R.,
Patterson.D., Reichwald.K., Rump.A., Schillhabel.M., Schudy.A.,
Zimmermann.W., Rosenthal.A., Kudoh.J., Shibuya.K., Kawasaki.K.,
Asakawa.S., Shintani.A., Sasaki.T., Nagamine.K., Mitsuyma.S.,
Antonarakis.S.E., Minoshima.S., Shimizu.N., Nordstiek.G.,
Hornischer.K., Brandt.P., Scharfe.M., Schoen.O., Desario.A.,
Reichelt.J., Kauer.G., Bloeker.H., Ramser.J., Beck.A., Klages.S.,
Hennig.S., Riesselmann.L., Dagand.E., Wehrmeyer.S., Borzym.K.,
Gardiner.K., Nizetic.D., Francis.F., Leirach.H., Reinhardt.R. and
Yaspo.M.I.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research

```

## COMMENT

Group \* Institute of Molecular Biotechnology, Genome Analysis \*  
 Keio University School of Medicine, Dept. of Molecular Biology \*  
 GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular  
 Genetics (addresses see below)  
 The Chromosome 21 Mapping and Sequencing Consortium consists of  
 \* RIKEN Genomic Sciences Center, Human Genome Research Group, \*  
 Sagami-hara 228-8555, Japan,  
 \* e-mail: sakaki@gsf.riken.go.jp  
 \* URL: http://hgp.gsc.riken.go.jp/  
 and  
 \* Institute of Molecular Biotechnology, Genome Analysis, \*  
 Beutenbergstrasse 11, D-07745 Jena, Germany,  
 \* e-mail: gscj-submit@genome.imb-jena.de  
 \* URL: http://genome.imb-jena.de/  
 and  
 \* Keio University School of Medicine, Dept. of Molecular Biology, \*  
 Tokyo 160-8582, Japan,  
 \* e-mail: shimizu@mb-med.keio.ac.jp  
 \* URL: http://adenine.dmb.med.keio.ac.jp/  
 and  
 \* GBF, Dept. of Genome Analysis,  
 \* Mascheroder Weg 1, D-38124 Braunschweig, Germany, \* e-mail:  
 info.genome@gbf.de  
 \* URL: http://genome.gbf.de/  
 and  
 \* Max-Planck Institute for Molecular Genetics,  
 \* Ihnestrasse 73, D-14195 Berlin, Germany,  
 \* e-mail: info-chr21@molgen.mpg.de  
 \* URL: http://chr21.rz-berlin.mpg.de/.

## FEATURES

## source

Location/Qualifiers  
 1..340000

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.3"

## source

-1..125946

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.3"

/clone="P160L9, 5' partial"  
 /clone\_lib="RPC11, 3-5 PAC library"  
 /note="Accession No. AF121897"

## source

73410..243533

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.3"

/clone="P70124"  
 /clone\_lib="RPC11, 3-5 PAC library"  
 /note="Accession No. AF064860"

## source

224137..2340000

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.3"

/clone="P206A10, 3' partial"  
 /clone\_lib="RPC11, 3-5 PAC library"  
 /note="Accession No. AF121782"

## source

286628..2340000

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.3"

/clone="BAC-291B3, 3' partial"  
 /clone\_lib="BAC library"  
 /note="Accession No. AF064857"

## repeat\_region

1088..1566

```

/note="L1M4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
repeat_region 2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region complement(2496..2878)
/note="TBI1C"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
repeat_region 2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region 2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region 3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
gene 3473..19157
/gene="SH3BGR"
Join(3473..3565,12649..12678,15331..15460,18739..19157)
mRNA /gene="SH3BGR"
CDS Join(3473..3565,12649..12678,15331..15426)
/partial
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CAB90445.1"
/db_xref="GI:7717370"
/translation="GSEKAEKGGTEAQKEGSDVGNLPEAQKNEBEGETATEETE
IAMEGAGEAEETETAEGERPGEDEDS"
exon 3473..3565
/gene="SH3BGR"
number=4
repeat_region 5542..5829
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region complement(5887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
repeat_region complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region 9131..9424
/note="AluSg"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region complement(10078..10429)
/note="TBI1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
repeat_region complement(10430..10651)

```

---

```

/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region 10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region 11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
exon 12649..12678
/gene="SH3BGR"
number=5
repeat_region complement(13115..13425)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region complement(13598..13892)
/note="AluY"

```

---

```

Query Match 25.7%; Score 318.6; DB 9; Length 340000;
Best Local Similarity 98.8%; Pred. No. 1.6e-64; Indels 0; Gaps 0;
Matches 321; Conservative 0; Mismatches 4;
QY 115 AGTTCTGGTCTGTTAAGTATAGAGAGCCGCCCAAGATGCAACAGTCCTGAAGGG 174
Db 268453 AGTTCTGGTCTGTTAAGTATAGAGAGCCGCCCAAGATGCAACAGTCCTGAAGGG 268512
QY 175 CTCCAGGCTCGTTCAACTGCACCGTCTCCAGGCTGGAAGCTCATCATGTGGCTCT 234
Db 268513 CTCCAGGCTCGTTCAACTGCACCGTCTCCAGGCTGGAAGCTCATCATGTGGCTCT 268572
QY 235 CAGTGACATGTTGTTGCTAAGGCTCAGGCCCATGAGCCCATCATCAATGACCGCTT 294
Db 268573 CAGTGACATGTTGTTGCTAAGGCTCAGGCCCATGAGCCCATCATCAATGACCGCTT 268632
QY 295 CACCTCTCAGAGTACGACCGGCGGAACTTCACCTCGGAGATGATCCCAATGT 354
Db 268633 CACCTCTCAGAGTACGACCGGCGGAACTTCACCTCGGAGATGATCCCAATGT 268692
QY 355 GGAGCCAGTGATTCGGGGAACATCAGATGCAAGCCTCCAGAACAGTCGCTGATGGATC 414
Db 268693 GGAGCCAGTGATTCGGGGAACATCAGATGCAAGCCTCCAGAACAGTCGCTGATGGATC 268752
QY 415 TGCTTACCTTACGTCCTCAAGTTATG 439
Db 268753 TGCTTACCTTACGTCCTCAAGTTATG 268777

```

---

```

RESULT 10
AC120145 149964 bp DNA linear HTG 09-MAR-2003
LOCUS Mus musculus clone RP24-216P24, WORKING DRAFT SEQUENCE, 9 unordered
DEFINITION pieces.
AC120145
AC120145.3 GI:28893697
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 149964)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP24-216P24

```



```

QY      292 CTTCACTCTCAGAGTACGACGCGGGGAACTTCACTCGAGATCATCATCA 351
Db      89388 CTTCACTATGCGAGTTACAAACAGACTGACAGCTTCATCTGGAGTTGATCATCATCA 89447

QY      352 TGTGGAGCCAGTCATTGCGGGGAACATCAGATCAGCCTCCAGAACAGTCGCTGCATGG 411
Db      89448 TGTGACGCGGAGTACTCGGGATCGCTGCAATGACGCTGCAGAACAGCCAGTGGTTGG 89507

QY      412 ATCTGCTTACCTTACGTCGTCGAAGT 436
Db      89508 ATCTGCTTCTCTCTCAGTCGCAAGT 89532

RESULT 11
AC120346 181510 bp DNA linear HTG 22-OCT-2002
LOCUS Mus musculus clone RP23-147E11, WORKING DRAFT SEQUENCE, 15 ordered
DEFINITION pieces.
AC120346
VERSION AC120346.4 GI:24211306
KEYWORDS HTG; HTGS PHASB2; HTGS DRAFT; HTGS_FULLTOR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 181510)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-147E11
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 181510)
Birren,B., Linton,B., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Lakocque,K., Lamazares,R.,
Lander,T., Lechoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Topham,K., Travers,M., Travis,N., Triggilio,J., Tesfaye,S., Theodore,J.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 181510)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Mathews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,

```

```

TITLE
JOURNAL
COMMENT
FEATURES
source

```

```

Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 22, 2002 this sequence version replaced gi:20514894.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20921
Center clone name: 147_E_11
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 177801 bases at least Q40
Consensus quality: 179340 bases at least Q30
Consensus quality: 179851 bases at least Q20
Insert size: 16000; agarose-fp
Quality coverage: 11.3 in Q20 bases; agarose-fp
Quality coverage: 11.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 23098: contig of 23098 bp in length
29099 29198: gap of 100 bp
29199 29864: contig of 666 bp in length
29865 29964: gap of 100 bp
29965 30244: contig of 280 bp in length
30245 30344: gap of 100 bp
30345 30971: contig of 627 bp in length
30972 31071: gap of 100 bp
31072 32648: contig of 1577 bp in length
32649 32748: gap of 100 bp
32749 37186: contig of 4438 bp in length
37187 37286: gap of 100 bp
37287 44243: contig of 6957 bp in length
44244 44343: gap of 100 bp
44344 54253: contig of 9910 bp in length
54254 54353: gap of 100 bp
54354 70476: contig of 16123 bp in length
70477 70576: gap of 100 bp
70577 86055: contig of 15479 bp in length
86056 86155: gap of 100 bp
86156 104095: contig of 17940 bp in length
104096 104195: gap of 100 bp
104196 123740: contig of 19545 bp in length
123741 123840: gap of 100 bp
123841 153195: contig of 29355 bp in length
153196 153295: gap of 100 bp
153296 180206: contig of 26911 bp in length
180207 180306: gap of 100 bp
180307 181510: contig of 1204 bp in length.
Location/Qualifiers
1. 181510
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-147E11"
/clone_lib="RPC1-23 Female Mouse BAC"

```

```
misc_feature 1. 29098
/note="assembly_fragment
clone_end:SP6
vector_side:left"
29199..29864
/note="assembly_fragment"
29965..30244
/note="assembly_fragment"
30345..30971
/note="assembly_fragment"
31072..32648
/note="assembly_fragment"
32749..37186
/note="assembly_fragment"
37287..44243
/note="assembly_fragment"
44344..54253
/note="assembly_fragment"
54354..70476
/note="assembly_fragment"
70577..86055
/note="assembly_fragment"
86156..104095
/note="assembly_fragment"
104196..123740
/note="assembly_fragment"
123841..153195
/note="assembly_fragment"
153296..180206
/note="assembly_fragment"
180307..181510
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 49236 a 40873 c 40708 g 49292 t 1401 others
ORIGIN
Query Match 15.1%; Score 187.4; DB 2; Length 181510;
Best Local Similarity 73.5%; Pred.No.1.9e-33;
Matches 239; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 112 TGTAGTTCTGGGTCTGTAATGAGTCATAGAGGGCCCCAGAAATGCAAGTCCTCGAA 171
Db 152360 TCTAGTTCGGATCCAGTTATCATAGATCATAGAGGCTCTCAGATGTACAGTCTTAA 152419
QY 172 GGGCTCCAGGTCGCTTCACTGACCGTCTCCAGGGCTGGAGCTCATCATCTGGGC 231
Db 152420 GGACTCAGAGGCTCACTTCACTGACCGTCACTGAGCTGGAAGCTTCTCATGTGGAC 152479
QY 232 TCTCAGTGACATGGTGGTGGTAAAGCTCAGGCCCCATGAGCCCATCATCAACCAATGACCG 291
Db 152480 TCTTAACCAATGGTGGTGGTGGTCTCACCACCAAGGACCCCATCATCAACCAACCG 152539
QY 292 CTTACCTCTAGAGTACGACGACGAGGCGGGAACCTTACCTCGGAGATGATCATCCACAA 351
Db 152540 CTTACCTCTAGAGTACGACGACGAGGCGGGAACCTTACCTCGGAGTGTATCATCCATGA 152599
QY 352 TGTGAGGCCAGTGTATCGGGGAACATCAGATCGAGCTCCAGAACAGTCCCTGCAATGG 411
Db 152600 TGTGAGGCCAGTGTATCGGGGAACATCAGATCGAGCTCCAGAACAGTCCCTGCAATGG 152659
QY 412 AFTCTGCTTACCTTACCGTCCAAAGTT 436
Db 152660 AFTCTGCTTCTCTCAGTGCAGGT 152684
```

```
RESULT 12
AC020851 AC020851 196900 bp DNA linear HTG 15-JUL-2000
LOCUS Mus musculus clone RP21-467L12, WORKING DRAFT SEQUENCE, 33
DEFINITION unorderd pieces.
ACCESSION AC020851
VERSION AC020851.2 GI:9211212
```

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
Mus musculus (house mouse)  
Mus musculus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 196900)  
DOE Joint Genome Institute.  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 196900)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 15, 2000 this sequence version replaced gi.6686457.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1437240  
Center clone name: RPCI-21\_467L12  
-----  
Summary Statistics  
Consensus quality: 164618 bases at least Q40  
Consensus quality: 180887 bases at least Q30  
Consensus quality: 183777 bases at least Q20  
Estimated insert size: 147000; pulse field gel estimation  
Estimated insert size: 193700; sum-of-contigs estimation  
Quality coverage: 5.67 in Q20 bases; pulse field gel estimation  
Quality coverage: 4.3 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1541: contig of 1541 bp in length  
1542 1641: gap of unknown length  
1642 3492: contig of 1851 bp in length  
3493 3592: gap of unknown length  
3593 4738: contig of 1146 bp in length  
4739 4838: gap of unknown length  
4839 7249: contig of 2410 bp in length  
7249 9179: gap of unknown length  
9179 9279: gap of unknown length  
9279 11493: contig of 2215 bp in length  
11494 11594: gap of unknown length  
11594 13608: contig of 2014 bp in length  
13608 13707: gap of unknown length  
13707 15946: contig of 2239 bp in length  
15946 16047: gap of unknown length  
16047 18392: contig of 2346 bp in length  
18392 18492: gap of unknown length  
18493 21331: contig of 2739 bp in length  
21331 21332: gap of unknown length  
21332 24662: contig of 3331 bp in length  
24662 24762: gap of unknown length  
24762 27140: contig of 2378 bp in length  
27140 27240: gap of unknown length  
27241 29950: contig of 2710 bp in length  
29951 30050: gap of unknown length  
30051 32245: contig of 2195 bp in length  
32245 32346: gap of unknown length  
32346 35688: contig of 3343 bp in length  
35688 35789: gap of unknown length  
35789 40387: contig of 4599 bp in length  
40388 40487: gap of unknown length  
40488 43590: contig of 3093 bp in length



AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.

TITLE B7-like molecules and uses thereof  
JOURNAL Patent: WO 0200710-A 11 03-JAN-2002;  
Amgen, Inc. (US)

## FEATURES

Location/Qualifiers  
1..895  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
53..865  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD28221.1"  
/db\_xref="GI:19575337"  
/translation="MEGSRDVLAVLILVLAQLTAGSSVQIIEGPQNVTVLKDESAHF  
NCTVHGKMLMTLNOMVLSLTQGPILTNFRFYASYNSTDSFISLIIHDVQPS  
DSGVSQSLQNSHGFSAFSLVSDSIEGEPALPTWAILLAVAFSLIILIIILII  
FCCCAARKEESTYQNEIRKSNMRTKADPETKLSKENYGVSSDEAKAAQTAS  
LPPKSAEVSUPEKRSSSLPYQLNKHQPGFATHPRVFSFDIASPQKVRNVTLV"  
BASE COUNT 250 a 223 c 212 g 210 t

## CDS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## REMARK

Query Match 15.0%; Score 185.6; DB 6; Length 895;  
Best Local Similarity 73.8%; Pred. No. 3.8e-33;  
Matches 236; Conservative 0; Mismatches 84; Indels 0; Gaps 0;  
QY 115 AGGTTCTGGTCTGGTAATGAAGTCATAGAAGGCCCCAGAGATCAACAGTCTCTGAAGG 174  
DB 109 AGCTTCGGATCCAGTTATCATGATCATAGAAGGCTCTCAGAACTTAACAGTCTCTAAAGGA 168  
QY 175 CTCGAGCTCGCTTCACTGACCGCTCCAGCGCTCCAGCGCTCCAGCGCTCCAGCGCTCT 234  
DB 169 CTCGAGCTCGCTTCACTGACCGCTCCAGCGCTCCAGCGCTCCAGCGCTCTCTCATGTGGACTCT 228  
QY 235 CAGTGACATGTTGTTGCTTAAGCTCAGCGCTCCAGCGCTCCAGCGCTCCAGCGCTCCAGCGCTT 294  
DB 229 TAACCAATGTTGTTGCTGAGTCTCACCACCAAGGACCCATCATCAACCAACAGCGCTT 288  
QY 295 CACCTCTCAGAGTACGACCGGCGGAACTTCACTCGGAGTATCATCCACATGT 354  
DB 289 CACCTATGCCAGTTACACACAGCTGACGATCATCTCGAGTTGATCATCCATGATG 348  
QY 355 CGAGCGCAGTATCGGGGAACATCAGATCAGCGCTCCAGAACAGTCCGCTGCGATGATC 414  
DB 349 CGAGCGCAGTATCGGGATCGGTGATCGAGCTCGAGACAGCAGCAGTGGTTGGATC 408  
QY 415 TGCTTACCTTACCGTCCAAAG 434  
DB 409 TGCCTCTCTCAGTGCAAG 428

## RESULT 15

LOCUS BC004806  
DEFINITION Mus musculus RIKEN cDNA 2010003D20 gene, mRNA (cDNA clone MGC:7960  
IMAGE:3584645), complete cds.

ACCESSION BC004806

VERSION BC004806.1 GI:13435932

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1556)

## REFERENCE

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Greenchenko, Y., Bouffard, G.G., Blakesley, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Dickson, M.C., Rodriguez, A.C., Touchman, J.W., Shreen, E.D., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)  
12477932  
2 (bases 1 to 1556)  
Strausberg, R.  
Direct Submission  
Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Place: 10 Row: m Column: 23  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

## FEATURES

## Location/Qualifiers

1..1556  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:7960 IMAGE:3584645"  
/tissue type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."  
/clone\_lib="NCI CGAP\_Mam1"  
/lab\_host="DH108"  
/note="Vector: pCMV-SPORT6"  
1..1556  
/gene="2010003D20Rik"  
/db\_xref="LocusID:72058"  
/db\_xref="MGI:1919308"  
55..867  
/codon\_start=1  
/product="RIKEN cDNA 2010003D20"  
/protein\_id="AA04806.1"  
/db\_xref="GI:13435933"  
/db\_xref="LocusID:72058"

## gene

## CDS

55..867  
/codon\_start=1  
/product="RIKEN cDNA 2010003D20"  
/protein\_id="AA04806.1"  
/db\_xref="GI:13435933"  
/db\_xref="LocusID:72058"

## BASE COUNT

447 a 364 c 367 g 378 t

## Query Match

15.0%; Score 185.6; DB 10; Length 1556;

Best Local Similarity		73.88;	Pred. No. 3.9e-33;		
Matches		236;	Conservative	0;	Mismatches 84;
				Indels	0;
				Gaps	0;
QY	115	AGGTTCTGGGCTCTGGTAATGAAGTCA	TAGAAGGCCCCAGATGCAACAGTCCCTGAAGGG	174	
Db	111	AGCTCCGGATCCAGTTATCAGATCAT	TAGAAGTCTCTCAGAAATGTACAGTCTTAAGGA	170	
QY	175	CTCCAGGCTCGCTTCAACTGACCGTCT	CCAGGGCTGGAAGCTCATCATGTGGGCTCT	234	
Db	171	CTCAGAGGCTCACTTCAACTGACCGTGA	CTCACGGCTGGAAGCTTCTCATGTGGACTCT	230	
QY	235	CAGTGACATGGTGGTGTCTAAGCGTCA	GGCCCATGGAGCCCATCATCAGCAATGACCGCTT	294	
Db	231	TAACCAATGGTGGTGTCTGAGTCTCA	CCCAAGGACCCATCATCAACCAACCGCTT	290	
QY	295	CACCTTCTCAGAGTAGCACCGGGCGG	GAACCTTCACCTCGGAGATGATCATCCACAATGT	354	
Db	291	CACCTATGCCAGTTACAACAGCACTG	ACAGCTTCATCTCGAGTTGATCATCCATGATGT	350	
QY	355	GGAGCCAGTGAATCGGGGAACATCAG	ATGCAGCCTCCAGAACAGTCGGCTGCATGGATC	414	
Db	351	GCAGCCAGTGAATCGGGATCCGTGCA	ATGCAGCCTGCAGAACAGGCCATGGGTTTGGATC	410	
QY	415	TGCTTACCTTACCGTCCAAG	434		
Db	411	TGCTTCTCTCAGTGCAG	430		

Search completed: November 12, 2003, 20:56:39  
Job time : 4799.54 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:39:35 ; Search time 366.843 Seconds

(without alignments)

9124.610 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240

Sequence: 1 aggtgtgagtcgacgaacaa.....gtaatacaactgtagtatatag 1240

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03.\*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1240	100.0	1240 24	ABK13030
2	1123.4	90.6	1175 24	ABK13028
3	1119	90.2	1168 24	ABK13029
4	1064	85.8	1139 24	ABK13031
5	777.6	62.7	1392 23	AAS92356
6	380.6	30.7	1195 24	ABK13032
7	188.8	15.2	401 22	AA136582
8	185.6	15.0	754 24	ABK13034

9	185.6	15.0	895	24	ABK13033	DNA encoding mouse
10	140.4	11.3	398	25	ABX55016	Bovine EST associa
11	53.4	4.3	277	21	ABK10239	Trinucleotide repe
12	52.8	4.3	7029	23	ABL29756	Drosophila melanog
13	51.8	4.2	389	21	ABK10248	Tri-nucleotide rep
14	51.8	4.2	397	20	AAK98991	Spinocerebellar at
15	51.8	4.2	403	21	ABK10240	Trinucleotide repe
16	49.6	4.0	742	22	AA195221	Human neuroblastom
17	49.4	4.0	46954	23	ABL16830	Drosophila melanog
18	48.4	3.9	3217	10	AAK91578	Rat androgen recep
19	48.4	3.9	3217	12	AAQ12002	Full-length rat an
20	48.4	3.9	4180	10	AAK91773	Rat androgen recep
21	47.6	3.8	2733	20	AAK84442	Mouse brain CNG-1
22	47.6	3.8	249999	25	ABZ80229	Human tramdorin ge
23	47.4	3.8	3894	23	ABL03353	Drosophila melanog
24	47.4	3.8	2341	23	ABL03352	Drosophila melanog
25	46.6	3.8	300	20	AAK98588	Human cancer cell
26	46.6	3.8	759	20	AAK98974	Human validated ca
27	46.6	3.8	2333	22	AAK75339	Human TGF-beta rec
28	46.6	3.8	14208	22	AAK78411	Human immune/haema
29	46.4	3.7	1983	23	ABL24483	Drosophila melanog
30	46.4	3.7	4044	23	ABL24482	Drosophila melanog
31	46.4	3.7	14152	22	AAS36073	Human cardiovascular
32	46.4	3.7	14152	22	AAK82884	Human immune/haema
33	46.2	3.7	381	23	ABV54466	Human prostate exp
34	46	3.7	513	23	ABV47976	Human prostate exp
35	46	3.7	3577	25	ABX16536	Canine cDNA encodi
36	45.8	3.7	405	23	ABV56338	Human prostate exp
37	45.8	3.7	2682	23	ABL16097	Drosophila melanog
38	45.8	3.7	8821	23	ABL16096	Drosophila melanog
39	45.6	3.7	1509	23	ABL24571	Drosophila melanog
40	45.6	3.7	3884	23	ABL24570	Drosophila melanog
41	45.4	3.7	578	22	AAS30941	Human cDNA encodin
42	45.4	3.7	3042	23	ABL28445	Drosophila melanog
43	45.4	3.7	3135	23	ABL25653	Drosophila melanog
44	45.4	3.7	5215	23	ABL28444	Drosophila melanog
45	45.4	3.7	5447	23	ABL25652	Drosophila melanog

#### ALIGNMENTS

##### RESULT 1

ID ABK13030 standard; cDNA; 1240 BP.

XX ABK13030;

XX 23-APR-2002 (first entry)

XX DNA encoding human B7-like protein, B7-L\_h3.

XX Human, B7-like protein; B7-L; antifertility; gynaecological;  
KW antitumour; cystostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 80..1240

XX FT /tag= a

XX FT /product= "B7-like protein, B7-L\_h3"

XX PN WQ200200710-A2.

XX PD 03-JAN-2002.

XX XX 28-JUN-2001; 2001WO-US20719.

```
XX 28-JUN-2000; 2000US-214512P.
PR 28-NOV-2000; 2000US-0729264.
XX (AMGE-) AMGEN INC.
XX
XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
XX WFI; 2002-130881/17.
XX P-PSDB; RAU75542.
XX
XX New B7-like polypeptides, polynucleotides and their modulators, useful
XX for diagnosing, preventing and treating reproductive, immune and
XX proliferative disorders, e.g. cancer and arteriosclerosis -
XX
XX Claim 1; Fig 3; 135pp; English.
XX
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).
XX The polypeptide, polynucleotide encoding it and antibody against (I) are
XX useful for treating B7-like polypeptide-related disease, disorders or
XX conditions including reproductive disorders (e.g. infertility,
XX miscarriage, preterm labour and delivery and endometriosis) and
XX proliferative disorders. Antibodies, soluble proteins comprising
XX extracellular domains and other regulators of B7-L polypeptides are
XX useful for enhancing the immune response to tumours. (I) plays a role in
XX growth and maintenance of cancer cells based on the observation of
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
XX polypeptide. Hence modulators of (I) are useful for the treatment of
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response
XX in allograft transplantation, graft versus host disease, T-cell
XX dependent B-cell mediated diseases and autoimmune diseases. B7-L
XX molecules are useful for alleviating the symptoms associated with
XX diseases involving chronic immune cell dysfunction or to treat
XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid
XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
XX purpura and psoriasis, chronic inflammatory disease such as
XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),
XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
XX are also useful as immunosuppressive agents for bone marrow and organ
XX transplantation or to prolong graft survival. B7-L molecules are also
XX useful for diagnosis and treatment of diseases involving abnormal cell
XX proliferation, including arteriosclerosis and vascular restenosis.
XX Antagonists of B7-L polypeptides are useful for alleviation of toxic
XX shock syndrome or allo sensitisation due to blood transfusions, and for
XX treatment of allergy, asthma and hypersensitivity reactions,
XX nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
XX pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
XX (extrinsic alveolitis) vasculopathies, coeliac disease, anaemia, and
XX thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
XX lymphoproliferative disorders such as multiple myeloma. The present
XX sequence represents the coding sequence of human B7-L h3.
XX
XX Sequence 1240 BP; 318 A; 319 C; 305 G; 298 T; 0 other;
XX
XX Query Match 100.0%; Score 1240; DB 24; Length 1240;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGTGTGAGTCCAGCCACAGTGTGGATCACTTCTTAGGTGCGCATAAACAAGCACCAT 60
XX
XX 1 AGGTGTGAGTCCAGCCACAGTGTGGATCACTTCTTAGGTGCGCATAAACAAGCACCAT 60
XX
XX 61 AACCTGGTGGCTTAGAAACAATGGAAAGGCATTCTCAGCGTTCCAGAGCTGTAGGTTT 120
XX
XX 61 AACCTGGTGGCTTAGAAACAATGGAAAGGCATTCTCAGCGTTCCAGAGCTGTAGGTTT 120
XX
XX 121 TGGGCTCTGATGAAGTATAGAGGCCGCCAGAAATGCAACAGTCTCTGAAGGCTCCCA 180
XX
XX 121 TGGGCTCTGATGAAGTATAGAGGCCGCCAGAAATGCAACAGTCTCTGAAGGCTCCCA 180
XX
XX 161 GGCTCGCTTCAACTGACCGCTCTCCAGGGCTGGAGCTCATCATGTGGGCTCTCAGTGA 240
XX
181 GGCTCGCTTCAACTGACCGCTCTCCAGGGCTGGAGCTCATCATGTGGGCTCTCAGTGA 240
241 CATGTGTGTCTAAAGCGTCAAGCCCATGGAGCCCATCATCAATGACCCCTTCACTC 300
241 CATGTGTGTCTAAAGCGTCAAGCCCATGGAGCCCATCATCAATGACCCCTTCACTC 300
301 TCAGAGGTACGACGAGGCGGGAACCTTCACTCCGAGATGATCATCCACATGTGGAGCC 360
301 TCAGAGGTACGACGAGGCGGGAACCTTCACTCCGAGATGATCATCCACATGTGGAGCC 360
361 CAGTCAATTCGGGGAAACATCAGATGAGCTCCAGAACAGTCCCTGATGATCTGCTTA 420
361 CAGTCAATTCGGGGAAACATCAGATGAGCTCCAGAACAGTCCCTGATGATCTGCTTA 420
421 CATTACCGTCCAGTATGAGGAGAGCTTCACTCCAGTGTATCTTGTAGTCCGCTGA 480
421 CATTACCGTCCAGTATGAGGAGAGCTTCACTCCAGTGTATCTTGTAGTCCGCTGA 480
481 GAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCCGGCTCCCGGATATTC 540
481 GAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCCGGCTCCCGGATATTC 540
541 CTGGAGCTCGGTCTCTGGTCAAGCTTCAAGCTTATTTTGTTCGGAGCCGAGCGA 600
541 CTGGAGCTCGGTCTCTGGTCAAGCTTCAAGCTTATTTTGTTCGGAGCCGAGCGA 600
601 CTTCAAGGTGAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 660
601 CTTCAAGGTGAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 660
661 CTTCAAGGTGAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 720
661 CTTCAAGGTGAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 720
721 TCGGTGTCTCCAGACACTGAGGCTGATTAATATCCAGTGTATTAATCAAGTTTACC 780
721 TCGGTGTCTCCAGACACTGAGGCTGATTAATATCCAGTGTATTAATCAAGTTTACC 780
781 GAGTTTAGGTTTTTCAATGCTTCTGGGCAAGCTTGGACTTGGACTTGGACTTGGACTT 840
781 GAGTTTAGGTTTTTCAATGCTTCTGGGCAAGCTTGGACTTGGACTTGGACTTGGACTT 840
841 GCTTCTGAGCGCGAGTGTACTCTTCAATAGCTGTCTGCTGCTGCTGCTGCTGCTGCTG 900
841 GCTTCTGAGCGCGAGTGTACTCTTCAATAGCTGTCTGCTGCTGCTGCTGCTGCTGCTG 900
901 TGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 TGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
961 TCAATTTCAAGAAATCTGAAAAAGAGAGAGCAAAACAAGAACTGAGACAGAAAGTGG 1020
961 TCAATTTCAAGAAATCTGAAAAAGAGAGAGCAAAACAAGAACTGAGACAGAAAGTGG 1020
1021 AAATGAAACTCCGGCTCAATTCAGATGCAACAAAGAACCCAGAAACCGCTTCTCTCC 1080
1021 AAATGAAACTCCGGCTCAATTCAGATGCAACAAAGAACCCAGAAACCGCTTCTCTCC 1080
1081 TCCAAATCTGTGAATCCAGTGTCTTCTGAAACAAGAAACAGTAGCTGTGGGCTCTCTCA 1140
1081 TCCAAATCTGTGAATCCAGTGTCTTCTGAAACAAGAAACAGTAGCTGTGGGCTCTCTCA 1140
1141 CCAGGGGCTGATCAACGTCACCCAGGCGCAGAGTGTATCCACAGGCTTCTTTTAATCT 1200
1141 CCAGGGGCTGATCAACGTCACCCAGGCGCAGAGTGTATCCACAGGCTTCTTTTAATCT 1200
1201 GGCCAGTCTCTGAGAGGTCAGTAAATACAACTGTAGTATAG 1240
1201 GGCCAGTCTCTGAGAGGTCAGTAAATACAACTGTAGTATAG 1240
```

RESULT 2

ABK13028  
 ID ABK13028 standard; cDNA; 1175 BP.  
 AC ABK13028;  
 DT 23-APR-2002 (first entry)  
 XX  
 XX DNA encoding human B7-like protein, B7-L\_h1.  
 XX Human; B7-like protein; B7-L; antiinfectivity; gynaecological;  
 KW antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
 KW antiasthmatic; nephrotropic; antibacterial; virucide; tumor; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 27..1175  
 XX /\*tag= a  
 XX /product= "B7-like protein, B7-L\_h1"  
 XX  
 XX WO2002007110-A2.  
 XX  
 XX 03-JAN-2002.  
 XX  
 XX 28-JUN-2001; 2001WO-US20719.  
 XX  
 XX 28-JUN-2000; 2000US-214512P.  
 XX 28-NOV-2000; 2000US-0729264.  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
 XX WPT; 2002-130881/17.  
 XX P-PSDB; AAU75540.  
 XX  
 XX New B7-like polypeptides, polynucleotides and their modulators, useful  
 XX for diagnosing, preventing and treating reproductive, immune and  
 XX proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX  
 XX Claim 1; Fig 1; 135pp; English.  
 XX  
 XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
 XX useful for treating B7-like polypeptide-related disease, disorders or  
 XX conditions including reproductive disorders (e.g. infertility,  
 XX miscarriage, preterm labour and delivery and endometriosis) and  
 XX proliferative disorders. Antibodies, soluble proteins comprising  
 XX extracellular domains and other regulators of B7-L polypeptides are  
 XX useful for enhancing the immune response to tumors. (I) plays a role in  
 XX growth and maintenance of cancer cells based on the observation of  
 XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 XX polypeptide. Hence modulators of (I) are useful for the treatment of  
 XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 XX in allograft transplantation, graft versus host disease, T-cell  
 XX dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 XX molecules are useful for alleviating the symptoms associated with  
 XX diseases involving chronic immune cell dysfunction or to treat  
 XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 XX purpura and psoriasis, chronic inflammatory disease such as  
 XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 XX are also useful as immunosuppressive agents for bone marrow and organ  
 XX transplantation or to prolong graft survival. B7-L molecules are also  
 XX useful for diagnosis and treatment of diseases involving abnormal cell

CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the coding sequence of human B7-L\_h1.  
 XX  
 XX SQ Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 other;  
 XX Query Match 90.6%; Score 1123.4; DB 24; Length 1175;  
 XX Best Local Similarity 99.9%; Pred. No. 1.8e-301;  
 XX Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 116 GGTTCCTGGTCTGGTAATGAAGTCATAGAAGCCGCCAGATGCAACAGTCTGAAGGCG 175  
 DB 51 GGTTCCTGGTCTGGTAATGAAGTCATAGAAGCCGCCAGATGCAACAGTCTGAAGGCG 110  
 QY 176 TCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGCGCTGGAAGCTCATCATGTGGGCTCTC 235  
 DB 111 TCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGCGCTGGAAGCTCATCATGTGGGCTCTC 170  
 QY 236 AGTGACATGGTGGTCTTAAGCGTCAAGGCCCATGAGGCCCATCATCAAAATGACCGTTC 295  
 DB 171 AGTGACATGGTGGTCTTAAGCGTCAAGGCCCATGAGGCCCATCATCAAAATGACCGTTC 230  
 QY 296 ACCTCTCAGAGGTAGCAGCCGCGGACCTTCACCTGGAGATGATCATCACCAATGTG 355  
 DB 231 ACCTCTCAGAGGTAGCAGCCGCGGACCTTCACCTGGAGATGATCATCACCAATGTG 290  
 QY 356 GAGCCAGTGAATTCGGGGAAACATCAGATGAGCCCTCCAGAACAGTCGGCTGATGGATCT 415  
 DB 291 GAGCCAGTGAATTCGGGGAAACATCAGATGAGCCCTCCAGAACAGTCGGCTGATGGATCT 350  
 QY 416 GCTTACCTTACCGTCCAGTTATGGAGAGCTGTTTCATCCAGTGTAACTCTTGTAGTC 475  
 DB 351 GCTTACCTTACCGTCCAGTTATGGAGAGCTGTTTCATCCAGTGTAACTCTTGTAGTC 410  
 QY 476 GCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCTCACCTGGACCCGGCTCCGGAT 535  
 DB 411 GCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCTCACCTGGACCCGGCTCCGGAT 470  
 QY 536 ATTTCTCGGAGCTCGTCTCTGTGTCAGCCATTCAAGCTATTATTTGTTCCGAGGCC 595  
 DB 471 ATTTCTCGGAGCTCGTCTCTGTGTCAGCCATTCAAGCTATTATTTGTTCCGAGGCC 530  
 QY 596 AGCGACCTTCAAAGTGCATGAGCATCTGGCTCTGACCCACAGACGAATGGGACTTGG 655  
 DB 531 AGCGACCTTCAAAGTGCATGAGCATCTGGCTCTGACCCACAGACGAATGGGACTTGG 590  
 QY 656 ACTTGGTGGCTTACCTGGAGAGCCCTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCACT 715  
 DB 591 ACTTGGTGGCTTACCTGGAGAGCCCTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCACT 650  
 QY 716 GTGATCTGGTGTCCCAACAGCATGGAGTGTATTAATATTCAGGTGTATTAATCAAGT 775  
 DB 651 GTGATCTGGTGTCCCAACAGCATGGAGTGTGTATTAATATTCAGGTGTATTAATCAAGT 710  
 QY 776 TTACCGAGTTTAGGTTTTTTCATTTGCTCTACTTGGGGCAAGTTGGACTTGGACTAGCAGGC 835  
 DB 711 TTACCGAGTTTAGGTTTTTTCATTTGCTCTACTTGGGGCAAGTTGGACTTGGACTAGCAGGC 770  
 QY 836 ACATGCTTCTGAGCCGAGTGTACTCTTAATAACGCTGTCTGTCTGCGCGCGTGTGT 895  
 DB 771 ACATGCTTCTGAGCCGAGTGTACTCTCTTAATAACGCTGTCTGTCTGCGCGCGTGTGT 830  
 QY 896 TGTGTGGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955  
 DB 831 TGTGTGGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890

QY 956 CGTATTCAATTTCAAAGAAATCTGAAAAGAGACAGACAAAGAACTGAGACAGAA 1015  
 DB 891 CGTATTCAATTTCAAAGAAATCTGAAAAGAGACAGACAAAGAACTGAGACAGAA 950  
 QY 1016 AGTGGAAATGAAAACCTCGGCTCAATTCAGATGAACAAAGACACAGAAACCGCTTCT 1075  
 DB 951 AGTGGAAATGAAAACCTCGGCTCAATTCAGATGAACAAAGACACAGACACCGCTTCT 1010  
 QY 1076 CTCCTCCCAATCTGTGAAATCCAGTATCTGAAACAAAGAACTGAGTGTGCGCTT 1135  
 DB 1011 CTCCTCCCAATCTGTGAAATCCAGTATCTGAAACAAAGAACTGAGTGTGCGCTT 1070  
 QY 1136 CCTCACCAGCGGGCTGATCAACGCTCCACCGAGCCAGCAAGTCAATCCACAGGCTTCTTTT 1195  
 DB 1071 CCTCACCAGCGGGCTGATCAACGCTCCACCGAGCCAGCAAGTCAATCCACAGGCTTCTTTT 1130  
 QY 1196 AATCGGCGAGTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1240  
 DB 1131 AATCGGCGAGTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1175

RESULT 3

ABK13029 ID ABK13029 standard; cDNA; 1168 BP.

AC ABK13029;

DT 23-APR-2002 (first entry)

DE DNA encoding human B7-like protein, B7-L<sub>h2</sub>.

KW Human; B7-like protein; B7-L; antiinfertility; gynaecological;  
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;  
 KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 8..1168

XX FT /\*tag= a

XX FT /product= "B7-like protein, B7-L<sub>h2</sub>"

XX WO200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.

XX 28-NOV-2000; 2000US-0729284.

XX (AMGE-) AMGEN INC.

XX Weiher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX WPI; 2002-130881/17.

XX P-PSDB; AAU75541.

XX New B7-like polypeptides, polynucleotides and their modulators, useful  
 XX for diagnosing, preventing and treating reproductive, immune and  
 XX proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX Claim 1; Fig 2; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
 XX useful for treating B7-like polypeptide-related disease, disorders or  
 XX conditions including reproductive disorders (e.g. infertility).

CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (I) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allo sensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemias,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the coding sequence of human B7-L<sub>h2</sub>.

XX Sequence 1168 BP; 299 A; 302 C; 288 G; 279 T; 0 other;

Query Match

Best Local Similarity 90.2%; Score 1119; DB 24; Length 1168;

Matches 1125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 106 AGAAGCTTAGGTTCTGGTCTGGTAAATGAAGTCAAGAGCCGCCAGAAATGCAACAGT 165  
 DB 34 AGACCCACCGGTTCTGGTCTGGTAAATGAAGTCAAGAGCCGCCAGAAATGCAACAGT 93  
 QY 166 CCGTAGGGGCTCCAGGCTCGTTCACCTGACCGTCTCCAGGGCTGGAGCTCATCAT 225  
 DB 94 CCGTAGGGGCTCCAGGCTCGTTCACCTGACCGTCTCCAGGGCTGGAGCTCATCAT 153  
 QY 226 GTGGGCTCTCAGTGACATGCTGGTGTAAAGCGTCAGGCCCATGGAGCCCATCACCAA 285  
 DB 154 GTGGGCTCTCAGTGACATGCTGGTGTAAAGCGTCAGGCCCATGGAGCCCATCACCAA 213  
 QY 286 TGACCGCTTCACCTCTCAGAGGTACGACCGGGGGGAACTTCCACCTCGGAGATGATCAT 345  
 DB 214 TGACCGCTTCACCTCTCAGAGGTACGACCGGGGGGAACTTCCACCTCGGAGATGATCAT 273  
 QY 346 CCACAATGTGGAGCCAGTGAATCGGGGAAACATCAGATGACGCTCCAGAACAGTCGCT 405  
 DB 274 CCACAATGTGGAGCCAGTGAATCGGGGAAACATCAGATGACGCTCCAGAACAGTCGCT 333  
 QY 406 GCATGGATCTGCTTACCTTACGTCCTCAAGTTATGGAGAGCTGTTCATCCAGTGTTAA 465  
 DB 334 GCATGGATCTGCTTACCTTACGTCCTCAAGTTATGGAGAGCTGTTCATCCAGTGTTAA 393  
 QY 466 TCTTGTAGTCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCCACATCGAGCCCG 525  
 DB 394 TCTTGTAGTCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCCACATCGAGCCCG 453  
 QY 526 GCTCCCGAATTTCTCGGGAGCTCGGTCTCTCGGTGAGCCATTCAGAGCTATTTATTTGT 585  
 DB 454 GCTCCCGAATTTCTCGGGAGCTCGGTCTCTCGGTGAGCCATTCAGAGCTATTTATTTGT 513  
 QY 586 TCCGAGCCCGAGCCAGCTTCAAGTGCAGTGAGATCTGGCTCTGACCCACAGAGCAA 645







QY 217 PDGTGGGGINPGVLSLPSLGLSLPTWKGVLGLAGTMLLTPTCTLTIRCCCRRCGCC 276  
 Db 58 KRQGGCGCCGCC-----GCCGCGGGGG--CGCCCRPRCCCRCCCTC 101  
 QY 277 --NCC-RCPCRCR 287  
 Db 102 CRTCCCTRCCTCCR 115

## RESULT 7

T15651  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 19-May-2000  
 C:Accession: I56551  
 R:Struyk, A.F.; Canoll, P.D.; Wolfgang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.  
 J. Neurosci. 15, 2141-2156, 1995  
 A:Title: Cloning of neurotrophin defines a new subfamily of differentially expressed neur  
 A:Reference number: I56551; MUID:95198094; PMID:7891157  
 A:Accession: I56551  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-344 <RES>  
 A:Cross-references: EMBL:U16945; NID:g755184; PID:AAA67445.1; PID:g755185  
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 5.9%; Score 122.5; DB 2; Length 188;  
 Best Local Similarity 61.3%; Pred. No. 0.057;  
 Matches 19; Conservative 1; Mismatches 8; Indels 3; Gaps 2;

QY 260 CTLTRCCCRRCGCC--NCC-RCPCRCR 287  
 Db 86 CCRPRCCCRRCCTCCCTCCCTCCCTCC 116

## RESULT 8

I56551  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 19-May-2000  
 C:Accession: I56551  
 R:Struyk, A.F.; Canoll, P.D.; Wolfgang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.  
 J. Neurosci. 15, 2141-2156, 1995  
 A:Title: Cloning of neurotrophin defines a new subfamily of differentially expressed neur  
 A:Reference number: I56551; MUID:95198094; PMID:7891157  
 A:Accession: I56551  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-344 <RES>  
 A:Cross-references: EMBL:U16945; NID:g755184; PID:AAA67445.1; PID:g755185  
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 5.8%; Score 121; DB 2; Length 344;  
 Best Local Similarity 24.1%; Pred. No. 0.14;  
 Matches 59; Conservative 37; Mismatches 93; Indels 56; Gaps 13;

QY 26 NARVLKGSQARFNCTVSGWKLIMVLSVRLPSPDLSQSAVILALTPQS 79

Db 44 NVTVROGESATRLCTDNRVTRVWMLNRSTLYAGNDKCLDPRVLLSN---TQTQY-- 98

QY 80 GGNFTSEMIHNVPEPSDGNIRCSLQ-----NSRLHGSAYLTQVMGELFIPS 133

Db 99 -----SIELQNVVDVDEGPTCSVQTDHFKTSRVH-----LIVQVSPKIVEISSDISIN 148

QY 134 ENPEPCBVTCL-----PGHWTLWLPDISWELGLLVSHSYFVPEPSDLSQSAVILALTPQS 189

Db 149 EGNNISLTATGRPE-----PVTWR---HISPKAVGFVSEDEYLE-----IQGITRQES 196

QY 190 GTLTCVATWKSLSKARKSATVNLTVIRCP-----QDTGGGINIPGVL-----SSLSLGSFL 240  
 Db 197 GEYECSSAS--NDVAPVVRVRYNTVNPFPYISEAKGTGVPVGKGTLOCEASAVPSAEFQ- 254  
 QY 241 PTWCK 245  
 Db 255 --WFK 257

## RESULT 9

T29757  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Dec-1999  
 C:Accession: T29757  
 R:Du, Z.; Le, T.T.; Wilson, R.  
 submitted to the EMBL Data Library, May 1997  
 A:Description: The sequence of C. elegans cosmid C09D1.  
 A:Reference number: Z20679  
 A:Accession: T29757  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-6642 <DUZ>  
 A:Cross-references: EMBL:AF003131; PID:AA54132.1; GSPDB:GN00019; CESP:unc-89  
 A:Experimental source: strain Bristol N2; clone C09D1  
 C:Genetics:  
 A:Gene: CESP:unc-89  
 A:Map position: 1  
 A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1;  
 /3; 5917/1; 6021/1; 6061/3; 6153/2; 6155/1; 6552/3; 6609/1

Query Match 5.7%; Score 119.5; DB 2; Length 6642;  
 Best Local Similarity 23.4%; Pred. No. 4.3;  
 Matches 51; Conservative 39; Mismatches 93; Indels 35; Gaps 8;

QY 20 VIEGPNARVLKGSQARFNCTVSGW--KLIMVLSVRLPSPDLSQSAVILALTPQS 77

Db 2077 VVDGPKSVTIKETETAEFKATIS-GFPAPTQKWTINEKIVEESTITITIKTEDVYT---- 2131

QY 78 DGGNFTSEMIHNVPEPSDGNIRCSLQNSRLHGSAYLTQVMGELFIPS-----VNLV 132

Db 2132 -----LKISNAKIEQTGTQKVTQNSAGQDSKQADLKVEPNVAPKFKSOLTDKVA 2182

QY 133 AENEPCVTCL---PSHWTLWLPDISWEL-GLLVSHSYFVPEPSDLSQSAVILALTPQS 188

Db 2183 DEGEPLRNWLELDGSPGT---EVSLLNGOPLTKSDTVQVVDGDTGYHYVIAAEDEM 2239

QY 189 NGTLTCVATWKSLSKARKSATVNLTL-----VIRCPQD 219

Db 2240 SGTLTAKAKNAAGCECTSAKVTGNGNKKKPEFVQAPQN 2277

## RESULT 10

JC4593  
 C:Species: Homo sapiens (man)  
 C:Date: 16-Apr-1996 #sequence\_revision 24-May-1996 #text\_change 24-Sep-1999  
 C:Accession: JC4593  
 R:Park, S.K.; Lee, H.S.; Lee, S.T.  
 J. Biochem. 119, 235-239, 1996  
 A:Title: Characterization of the human full-length PK7 cDNA encoding a receptor prote  
 A:Reference number: JC4593; MUID:97037064; PMID:8882711  
 A:Accession: JC4593  
 A:Molecule type: mRNA  
 A:Residues: 1-1070 <PAR>  
 A:Cross-references: GB:U40271; NID:g1322231; PID:AA50484.1; PID:g1322232  
 C:Comment: This protein is a member of receptor protein tyrosine kinase family, but pr

C:Genetics:  
 A:Gene: GDB:PK7  
 A:Cross-references: GDB:134760; OMIM:601890  
 A:Map position: 6p21.1-6p12.2  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo

[illegible]

Db 145 TPEVKMGKVDGDIRDAAIIDVTSSFTS-LVLNDVNRDYSKGYTLTLNSSGTSKSAFVTVR 203  
 QY 119 VMGELFIPSNVLVAENPECEVTCLPSHWTLPLDI----- 153  
 Db 204 VL-DTPSPVNLKVTETLKDSVSI-----TWEPFLDGGSKIKNYIVEKREARKSYAAV 257  
 QY 154 -----SWELGLLVSHSSYFVPEPSDLQSAVSILALT-----PQNSGFLTCVAT 197  
 Db 258 VTNCHKNSWKIDQLQEGCSYF-RVTAENEXYIGLAARTADPIKVAEVPQPGKITVDVV 316  
 QY 198 WKSLKARKSATVNLTVIRCPQDTGG 222  
 Db 317 -----TRNSVLSWTK---PEHDDG 333  
 RESULT 14  
 I49583  
 differentiation antigen - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C/Accession: I49583  
 R/Law, C.  
 J. Immunol. 151, 175-187, 1993  
 A/Title: Organization of the murine Cd22 locus. Mapping to chromosome 7 and characterization  
 A/Reference number: I49583; MUID:93315634; PMID:8100843  
 A/Accession: I49583  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-862 <RES>  
 A/Cross-references: GB:I49583; NID:g348965; PIDN:AAA02562.1; PID:g348966  
 C/Genetics:  
 A/Gene: CD22  
 Query Match 5.5%; Score 114.5; DB 2; Length 862;  
 Best Local Similarity 19.1%; Pred. No. 1.1;  
 Matches 86; Conservative 36; Mismatches 126; Indels 203; Gaps 18;  
 QY 10 DPPGGSGNVEIQPNARVLKSGQAFN-----VLKPGVLRQKVTWDSMPVSCAACHKCSWALPVLNVHYA 519  
 Db 472 NPOGSGS-----RYDQGNFTSE---MIHNVPS 95  
 QY 52 LSDMWLVSRPMEPIITNDRFTSQ-----FDSSGQWKL-----TWAA 51  
 Db 520 PRDVKVLKVPASIRAGQVLLQCDFAESNPAEVRFWKNGSLVQEGYLSFGVSPE 579  
 QY 96 DSGNIRSLQNSRLHGSAYLTQVWNGELFIPSNLV-----VAENEP 138  
 Db 580 DSGNYNCMVNS-----IGETLSQAMNLQVLYAPRRLRVSIISPGDHVMEGKA 627  
 QY 139 EVTCLP-----SHWTWLPDISWELGLLVSHSSYFVPEPSDLQSAVSILALT-----QSN 189  
 Db 628 TSCSDANPPIQYTW-----FDSSGQWKL-----TWAA 51  
 QY 190 GTTCTVATWKSARKSATVNLTVIRCPQDTGGINIPGVLSPLSLPTWKGVLG 249  
 Db 670 GSVRCGT-NGIGTGESPPSLTVVYSPETIG-----KRVAG 706  
 QY 250 LAGTMLTPTCTLTIRCCCRCCGCCNCCRCFCRKGFRQ-----FK 298  
 Db 707 LGFLCTI---CILAI-----WGMKIQKWKQNSRQOGLQE 738  
 QY 299 KSE-----KEKTKETETESGNSG-----YNSDQKTTDT 330  
 Db 739 NSSQCSFFVRNKKARPTLESGPQSCYNPAMDDTVSYALLRPESDMHAGDGPAT 798  
 QY 331 ASLPKSCS-----SDPEQRNSC 350  
 Db 799 QAPPNNNSDVSIVIQKPMGDYENVNPS 829  
 RESULT 15  
 I38344  
 titin, cardiac muscle [validated] - human

N/Alternate names: connectin  
 C/Contains: serine/threonine-specific protein kinase (EC 2.7.1.1-)  
 C/Species: Homo sapiens (man)  
 C/Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 15-Sep-2000  
 C/Accession: I38344; I38345; S20897; S20898; S20899; S63665; S37393  
 R/Labeit, S.; Kolmerer, B.  
 Science 270, 293-296, 1995  
 A/Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.  
 A/Reference number: A57430; MUID:96026330; PMID:7569978  
 A/Accession: I38344  
 A/Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL  
 A/Molecule type: mRNA  
 A/Residues: 1-26926 <LAB1>  
 A/Cross-references: EMBL:X90568; NID:g1017424; PID:g1017425  
 R/Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.  
 Biochemistry 34, 553-561, 1995  
 A/Title: Dissecting titin into its structural motifs: identification of an alpha-helix  
 A/Reference number: I38345; MUID:95119041; PMID:7819249  
 A/Accession: I38345  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1577-2014 <MUS>  
 A/Cross-references: EMBL:X83270; NID:g602579; PIDN:CAA58243.1; PID:g602580  
 A/Note: Conformation and properties are reported for a synthetic peptide corresponding  
 R/Labeit, S.; Gautel, M.; Lakey, A.; Trinick, J.  
 EMBO J. 11, 1711-1716, 1992  
 A/Title: Towards a molecular understanding of titin.  
 A/Reference number: S20897; MUID:92258380; PMID:1582406  
 A/Accession: S20898  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 13597-14200, 'I', 14202-14696 <LAB2>  
 A/Cross-references: EMBL:X64698; NID:g37192; PIDN:CAA45939.1; PID:g37193  
 A/Accession: S20897  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 16330-16382, 'S', 16384-16756, 'F', 16758-16860 <LAB3>  
 A/Cross-references: EMBL:X64699; NID:g37190; PIDN:CAA45940.1; PID:g37191  
 A/Accession: S20899  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 16729-26825 <LAB4>  
 A/Cross-references: EMBL:X92412; NID:g1236761  
 R/Gautel, M.; Leonard, K.; Labeit, S.  
 EMBO J. 12, 3827-3834, 1993  
 A/Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differenti  
 A/Reference number: S37393; MUID:94008990; PMID:8404852  
 A/Accession: S37393  
 A/Molecule type: mRNA  
 A/Residues: 26831-26926 <GAU>  
 R/Imrota, S.; Politou, A.S.; Pastore, A.  
 submitted to the Brookhaven Protein Data Bank, February 1996  
 A/Reference number: A66736; PDB:1TIT  
 A/Contents: annotation; conformation by (1)H-NMR, residues 5253-5341  
 R/Pfuhl, M.; Pastore, A.  
 submitted to the Brookhaven Protein Data Bank, August 1996  
 A/Reference number: A66201; PDB:1NCT  
 A/Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155  
 C/Genetics:  
 A/Gene: GDB:TTN  
 A/Cross-references: GDB:I27867; OMIM:188840  
 A/Map position: 2q31-q32  
 C/Function:  
 A/Description: structural protein forming filaments in striated muscle

C; Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro C; Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco structural protein

F;24752-25008/Domain: protein kinase homology <KIN>

F; 84, 177, 905, 2276, 2378, 2459, 2481, 2563, 2669, 2763, 2896, 3088, 3179, 3384, 3432, 3628, 3772, 4068.

98, 11066, 11488, 11515, 11635, 11949, 12170, 12478, 12526, 12645, 12875, 13001, 13036, 13295, 13540, 13771, 13999, 14000, 14001, 14002, 14003, 14004, 14005, 14006, 14007, 14008, 14009, 14010, 14011, 14012, 14013, 14014, 14015, 14016, 14017, 14018, 14019, 14020, 14021, 14022, 14023, 14024, 14025, 14026, 14027, 14028, 14029, 14030, 14031, 14032, 14033, 14034, 14035, 14036, 14037, 14038, 14039, 14040, 14041, 14042, 14043, 14044, 14045, 14046, 14047, 14048, 14049, 14050, 14051, 14052, 14053, 14054, 14055, 14056, 14057, 14058, 14059, 14060, 14061, 14062, 14063, 14064, 14065, 14066, 14067, 14068, 14069, 14070, 14071, 14072, 14073, 14074, 14075, 14076, 14077, 14078, 14079, 14080, 14081, 14082, 14083, 14084, 14085, 14086, 14087, 14088, 14089, 14090, 14091, 14092, 14093, 14094, 14095, 14096, 14097, 14098, 14099, 14100, 14101, 14102, 14103, 14104, 14105, 14106, 14107, 14108, 14109, 14110, 14111, 14112, 14113, 14114, 14115, 14116, 14117, 14118, 14119, 14120, 14121, 14122, 14123, 14124, 14125, 14126, 14127, 14128, 14129, 14130, 14131, 14132, 14133, 14134, 14135, 14136, 14137, 14138, 14139, 14140, 14141, 14142, 14143, 14144, 14145, 14146, 14147, 14148, 14149, 14150, 14151, 14152, 14153, 14154, 14155, 14156, 14157, 14158, 14159, 14160, 14161, 14162, 14163, 14164, 14165, 14166, 14167, 14168, 14169, 14170, 14171, 14172, 14173, 14174, 14175, 14176, 14177, 14178, 14179, 14180, 14181, 14182, 14183, 14184, 14185, 14186, 14187, 14188, 14189, 14190, 14191, 14192, 14193, 14194, 14195, 14196, 14197, 14198, 14199, 14200, 14201, 14202, 14203, 14204, 14205, 14206, 14207, 14208, 14209, 14210, 14211, 14212, 14213, 14214, 14215, 14216, 14217, 14218, 14219, 14220, 14221, 14222, 14223, 14224, 14225, 14226, 14227, 14228, 14229, 14230, 14231, 14232, 14233, 14234, 14235, 14236, 14237, 14238, 14239, 14240, 14241, 14242, 14243, 14244, 14245, 14246, 14247, 14248, 14249, 14250, 14251, 14252, 14253, 14254, 14255, 14256, 14257, 14258, 14259, 14260, 14261, 14262, 14263, 14264, 14265, 14266, 14267, 14268, 14269, 14270, 14271, 14272, 14273, 14274, 14275, 14276, 14277, 14278, 14279, 14280, 14281, 14282, 14283, 14284, 14285, 14286, 14287, 14288, 14289, 14290, 14291, 14292, 14293, 14294, 14295, 14296, 14297, 14298, 14299, 14300, 14301, 14302, 14303, 14304, 14305, 14306, 14307, 14308, 14309, 14310, 14311, 14312, 14313, 14314, 14315, 14316, 14317, 14318, 14319, 14320, 14321, 14322, 14323, 14324, 14325, 14326, 14327, 14328, 14329, 14330, 14331, 14332, 14333, 14334, 14335, 14336, 14337, 14338, 14339, 14340, 14341, 14342, 14343, 14344, 14345, 14346, 14347, 14348, 14349, 14350, 14351, 14352, 14353, 14354, 14355, 14356, 14357, 14358, 14359, 14360, 14361, 14362, 14363, 14364, 14365, 14366, 14367, 14368, 14369, 14370, 14371, 14372, 14373, 14374, 14375, 14376, 14377, 14378, 14379, 14380, 14381, 14382, 14383, 14384, 14385, 14386, 14387, 14388, 14389, 14390, 14391, 14392, 14393, 14394, 14395, 14396, 14397, 14398, 14399, 14400, 14401, 14402, 14403, 14404, 14405, 14406, 14407, 14408, 14409, 14410, 14411, 14412, 14413, 14414, 14415, 14416, 14417, 14418, 14419, 14420, 14421, 14422, 14423, 14424, 14425, 14426, 14427, 14428, 14429, 14430, 14431, 14432, 14433, 14434, 14435, 14436, 14437, 14438, 14439, 14440, 14441, 14442, 14443, 14444, 14445, 14446, 14447, 14448, 14449, 14450, 14451, 14452, 14453, 14454, 14455, 14456, 14457, 14458, 14459, 14460, 14461, 14462, 14463, 14464, 14465, 14466, 14467, 14468, 14469, 14470, 14471, 14472, 14473, 14474, 14475, 14476, 14477, 14478, 14479, 14480, 14481, 14482, 14483, 14484, 14485, 14486, 14487, 14488, 14489, 14490, 14491, 14492, 14493, 14494, 14495, 14496, 14497, 14498, 14499, 14500, 14501, 14502, 14503, 14504, 14505, 14506, 14507, 14508, 14509, 14510, 14511, 14512, 14513, 14514, 14515, 14516, 14517, 14518, 14519, 14520, 14521, 14522, 14523, 14524, 14525, 14526, 14527, 14528, 14529, 14530, 14531, 14532, 14533, 14534, 14535, 14536, 14537, 14538, 14539, 14540, 14541, 14542, 14543, 14544, 14545, 14546, 14547, 14548, 14549, 14550, 14551, 14552, 14553, 14554, 14555, 14556, 14557, 14558, 14559, 14560, 14561, 14562, 14563, 14564, 14565, 14566, 14567, 14

[illegible]

F:16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18

,21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248

```
F;26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental
```

Query Match 5.5%; Score 114.5; DB 1; Length 26926;

Best Local Similarity 22.1%; Pred. No. 45;

Matches 62; Conservative 31; Mismatches 87; Indels 101; Gaps 12;

Qy 3 AGAMENRDPGSGGNEVIEGPQ-----NARV-----LKGSQARFNCTVS 42

[illegible]

QY 43 QGWKLIMWALSDMVVLSVRPMEPIITNDRFTSORYDOGGNFTSEMIHNVPSDSGNIRC 102

QZ IS COMBINATION OF VARIOUS KEYWORDS IN ORDER TO IDENTIFY THE  
RELEVANT INFORMATION AND PROVIDE A SUMMARY OF THE RESULTS.

07 103 STONSP1HGSAVI.TVOVMGEI.F1RPSVNI.VAENEDCEHTAT BCUHTWTI BDT 153

QY 103 STQNSRTHGSAYLTIVQVMGELFIPSVNLVVAENEPCEVTCLPSSHWTWLPDI----- 153

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

```

QY 154 -----SWELGLLVSHSSYF-----VPEPSDLQSAVSI 181

```

.....

QY 182 LALTPQNGTLTCVATWWSLKARKSATVNLTVIRCPQDTGG 222

Search completed: November 13, 2003, 03:34:28

Job time : 17.3865 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 23:23:44 ; Search time 9.36568 Seconds  
(without alignments)  
1938.172 Million cell updates/sec

Title: US-09-729-264-4

Perfect score: 2088

Sequence: 1 MWAGAMENRDPGSGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	145.5	7.0	404	1	RAGE_HUMAN
2	136	6.5	1447	1	DCC_HUMAN
3	132	6.3	1447	1	DCC_MOUSE
4	126	6.0	1914	1	KWLS_HUMAN
5	125	6.0	337	1	GSSA_CHICK
6	122	5.8	353	1	CEPU_CHICK
7	121	5.8	344	1	NTRI_RAT
8	119.5	5.7	6632	1	UN89_CABEL
9	119	5.7	1070	1	PTK7_HUMAN
10	115	5.5	344	1	NTRI_HUMAN
11	115	5.5	344	1	NTRI_MOUSE
12	114.5	5.5	862	1	CD22_MOUSE
13	113	5.4	569	1	SILF_MOUSE
14	113	5.4	1461	1	NEO1_HUMAN
15	112.5	5.4	345	1	OPCM_HUMAN
16	112.5	5.4	416	1	RAGE_BOVIN
17	111.5	5.3	345	1	OPCM_BOVIN
18	111.5	5.3	1443	1	NEO1_CHICK
19	111	5.3	620	1	SMP_COTUJ
20	111	5.3	1377	1	NEO1_RAT
21	110.5	5.3	345	1	OPCM_RAT
22	110.5	5.3	4391	1	PGBM_HUMAN
23	109.5	5.2	249	1	CSP_DROME
24	109.5	5.2	364	1	CD33_HUMAN
25	109	5.2	3375	1	UN52_CABEL
26	106.5	5.1	524	1	BUTY_MOUSE
27	106	5.1	1092	1	NCA2_XENLA
28	104	5.0	319	1	A33_HUMAN
29	104	5.0	764	1	ICCR_DROME
30	103.5	5.0	365	1	ICCR_MOUSE
31	103.5	5.0	3707	1	PGBM_MOUSE
32	102.5	4.9	333	1	AMAL_DROME
33	102.5	4.9	442	1	SIL6_HUMAN

#### RESULT 1

ID	RAGE_HUMAN	STANDARD;	PRT;	404 AA.
AC	Q15109; Q15279; Q9H2X7; Q9Y3R3;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Advanced glycosylation and product-specific receptor precursor			
DE	(Receptor for advanced glycosylation end products)			
GN	AGER OR RAGE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1].			
RC	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	TISSUE=Lung;			
RX	MEDLINE=92340547; PubMed=1378843;			
RA	Neeper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,			
RA	Elliston K., Stern D., Shaw A.;			
RT	"Cloning and expression of a cell surface receptor for advanced			
RT	glycosylation end products of proteins.";			
RL	J. Biol. Chem. 267:14998-15004(1992).			
RN	[2].			
RC	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	MEDLINE=95137587; PubMed=7835890;			
RA	Sugaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,			
RA	Inoko H., Ikemura T.;			
RT	"Three genes in the human MHC class III region near the junction with			
RT	the class II: gene for receptor of advanced glycosylation end			
RT	products, PEX2 homeobox gene and a notch homolog, human counterpart			
RT	of mouse mammary tumor gene int-3.";			
RL	Genomics 23:408-419(1994).			
RN	[3].			
RC	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	Rowen L., Danks C., Baskin D., Faust J., Loretz C., Ahearn M.E.,			
RA	Banta A., Spies T., Hood L.;			
RT	"Sequence determination of 300 kilobases of the human class III MHC			
RT	locus.";			
RL	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.			
RN	[4].			
RC	SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.			
RA	Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,			
RA	Yamamoto H.;			
RT	"Molecular heterogeneity of the receptor for advanced glycation			
RT	endproducts.";			
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5].			
RC	SEQUENCE FROM N.A. (ISOFORM 2).			
RA	Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,			
RA	Schuler A., Huber G.;			
RT	"cDNA cloning of a novel secreted isoform of the human receptor for			
RT	advanced glycation end products (RAGE) and characterization of cells			
RT	co-expressing cell-surface scavenger receptors and Swedish mutant			
RT	amyloid precursor protein.";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			

Q62786 rattus norv  
P06418 homo sapien  
P78310 homo sapien  
Q9wv91 mus musculu  
Q8z0j8 rattus norv  
Q24372 drosophila  
Q14162 homo sapien  
Q9p0x4 homo sapien  
P36774 rattus norv  
P34082 drosophila  
P14781 gallus gall  
P33470 porcine tra

[6]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Lung;  
 RX MEDLINE=2338257; PubMed=12477932;  
 RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [7]  
 RP SEQUENCE OF 1-12 FROM N.A.  
 RA Hudson B.I., Futers T.S.;  
 RT "Novel polymorphisms in the receptor for advanced glycation  
 end-products (RAGE) gene.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Mediates interactions of advanced glycosylation end  
 CC products (AGE). These are nonenzymatically glycosylated proteins  
 CC which accumulate in vascular tissue in aging and at an accelerated  
 CC rate in diabetes. Receptor for amyloid beta peptide.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).  
 CC Secreted (isoform 2).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC Name=2; Synonyms=RAGESEC;  
 CC IsoId=Q15109-2; Sequences=VSP\_002551, VSP\_002552;  
 CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M91211; AAA03574.1; -  
 CC DR EMBL; D28769; BAA05958.1; -  
 CC DR EMBL; U93336; ABA47491.1; -  
 CC DR EMBL; AB036432; BAA89369.1; -  
 CC DR EMBL; AJ134822; CAB43108.1; -  
 CC DR EMBL; BC020669; AAB20669.1; -  
 CC DR EMBL; AF208289; AAG35728.1; -  
 CC DR PIR; I61596; I61596.  
 CC Genew; HGNC:320; AGER.  
 CC MIM; 600214; -  
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
 CC DR GO; GO:004888; F:transmembrane receptor activity; TAS.  
 CC DR GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; TAS.  
 CC DR GO; GO:006954; P:inflammatory response; TAS.  
 CC DR InterPro; IPR007110; Ig-like.  
 CC DR InterPro; IPR003598; Ig\_c2.  
 CC DR InterPro; IPR003006; Ig\_MHC.  
 CC DR Pfam; PF00047; Ig; 2.  
 CC DR SMART; SM00408; Igc2; 1.

DR PROSITE; PS50835; IG LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;  
 FT SIGNAL 1 22  
 FT CHAIN 23 404  
 FT POTENTIAL  
 FT ADVANCED GLYCOSYLATION END PRODUCT-  
 FT SPECIFIC RECEPTOR.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT IG-LIKE V-TYPE.  
 FT IG-LIKE C2-TYPE 1.  
 FT IG-LIKE C2-TYPE 2.  
 FT POTENTIAL.  
 FT DISULFID 38 99  
 FT DISULFID 144 208  
 FT DISULFID 259 301  
 FT CARBOHYD 25 25  
 FT CARBOHYD 81 81  
 FT CARBOHYD 380 384  
 FT VARSPLIC 54 67  
 FT MISSING (in isoform 2).  
 FT /FTID=VSP\_002551.  
 FT /FTID=VAR\_011338  
 FT /FTID=VAR\_011338  
 FT M -> G (IN REF. 1).  
 FT CONFLICT 1 1  
 FT SEQUENCE 404 AA; 42802 MW; 0D584C436C30CE7 CRC64;  
 SQ  
 Query Match 7.0%; Score 145.5; DB 1; Length 404;  
 Best Local Similarity 23.5%; Pred. No. 0.00025;  
 Matches 77; Conservative 34; Mismatches 94; Indels 123; Gaps 15;  
 QY 64 EPIITNDRETS-----QRYDQGGNFT--SEMIHNVPESDSGNIR----CSLQNSLHGS 112  
 DB 162 KPLVPNEKGVSVKEQTRRHPTGLFTLQSELM--VTPARGDPRPTFSCSFSPGLPRHR 218  
 QY 113 AYLTVQVMGELFIP---SVNLVVAENP-----CEVTCPLPSHWTLWLPDLS 154  
 DB 219 ALRTAPIQPRVWPEVPLEVQLVW---BPEGGAVPGGVTVTLTCEVPAQFS-----PQIH 270  
 QY 155 WELGLLVSHSSYFVPEPSDLSQSAVSILALTQPSNGTLTCVATWKSLLKARKSATVNLTVI 214  
 DB 271 WMKD-----GVPPLPPSPVLILPIGPDQGTYSVATHSSHGQESRAVSISII 321  
 QY 215 RCPDGTGGGINIPGVLSLPSLPSLPTWGVGLGLAGTMLLT-----PTCTLTIRCC 267  
 DB 322 E-PGEEG-----PTAGSVGGSLGTALALGLTGLGTAALLIGVI 361  
 QY 268 CCRRCGCGNCCRCFCRCRRKGRFRIQFKSEKTEK--NKETETESGNSGNSGNSDQ 325  
 DB 362 LMQRR-----ORGEERKAPENQEBEERAEIN----- 389  
 QY 326 KTTDTASLPKSCSSDPEQRNSSCGPP 353  
 DB 390 -----QSEEPAGESSTGGP 404  
 RESULT 2  
 DCC HUMAN  
 ID DCC HUMAN STANDARD; PRT; 1447 AA.  
 AC P43146;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Tumor suppressor protein DCC precursor (Colorectal cancer suppressor).  
 GN DCC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95011532; PubMed=7926722;  
 RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,  
 RA Vogelstein B.,  
 RT "The DCC gene product in cellular differentiation and colorectal  
 RT tumorigenesis.";  
 RL Genes Dev. 8:1174-1183(1994).  
 RN [2]  
 RP SEQUENCE OF 1-750 FROM N.A.  
 RX MEDLINE=90100559; PubMed=2294591;  
 RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,  
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Vogelstein B.,  
 RA Vogelstein B.,  
 RT "Identification of a chromosome 18q gene that is altered in  
 RT colorectal cancers.";  
 RL Science 247:49-56(1990).  
 RN [3]  
 RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBELD EXONS).  
 RX MEDLINE=9121517; PubMed=1991322;  
 RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,  
 RA Oliner J.D., Kinzler K.W., Vogelstein B.,  
 RT "Scrambled exons.";  
 RL Cell 64:607-613(1991).  
 RN [4]  
 RP GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.  
 RX MEDLINE=94245241; PubMed=8188295;  
 RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,  
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.,  
 RT "The DCC gene: structural analysis and mutations in colorectal  
 RT carcinomas.";  
 RL Genomics 19:525-531(1994).  
 RN [5]  
 RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.  
 RX MEDLINE=94243823; PubMed=8187090;  
 RA Miyake S., Nagai K., Yoshino K., Oco M., Endo M., Yuasa Y.,  
 RT "Point mutations and allelic deletion of tumor suppressor gene DCC in  
 RT human esophageal squamous cell carcinomas and their relation to  
 RT metastasis.";  
 RL Cancer Res. 54:3007-3010(1994).  
 CC -!- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL  
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.  
 CC -!- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO  
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC  
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR  
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS  
 CC METASTASIS OF ESOPHAGEAL SQUAMOUS CELL CARCINOMAS.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.  
 CC -!- SIMILARITY: Contains 6 fibronectin type III domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M63698; AAA52181.1; -.  
 DR PIR; A54100; A54100.  
 DR HSSP; P56276; ITLK.  
 DR Genew; HGNC:2701; DCC.  
 DR MIM; 120470; -.  
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO; GO:0007409; P:axogenesis; TAS.  
 DR GO; GO:0006917; P:induction of apoptosis; TAS.  
 DR GO; GO:0007048; P:oncogenesis; TAS.  
 DR InterPro; IPR003961; FN III.  
 DR InterPro; IPR003962; FNIII subd.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig-C2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00041; fn3; 6.  
 DR Pfam; PF00047; ig; 4.  
 DR PRINTS; PR00014; FNYPEIII.  
 DR SMART; SM00060; FN3; 6.  
 DR SMART; SM00408; Igc2; 3.  
 DR PROSITE; PS00835; IG LIKE; 4.  
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;  
 KW Repeat; Anti-oncogene; Disease mutation; Polymorphism.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 1447 TUMOR SUPPRESSOR PROTEIN DCC.  
 FT DOMAIN 1098 1122 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1123 1447 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 26 135 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 139 229 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 234 326 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 331 416 IG-LIKE C2-TYPE 4.  
 FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 525 618 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 619 716 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 722 816 FIBRONECTIN TYPE-III 4.  
 FT DOMAIN 840 940 FIBRONECTIN TYPE-III 5.  
 FT DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.  
 FT DISULFID 61 117 BY SIMILARITY.  
 FT DISULFID 161 212 BY SIMILARITY.  
 FT DISULFID 261 310 BY SIMILARITY.  
 FT DISULFID 352 400 BY SIMILARITY.  
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 702 702 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 168 168 M > T (in esophageal carcinoma).  
 FT VARIANT 201 201 /FTID=VAR 003909.  
 FT VARIANT 201 201 R -> G.  
 FT VARIANT 1375 1375 /FTID=VAR 003910.  
 FT VARIANT 1375 1375 P -> H (in a colorectal carcinoma).  
 FT CONFLICT 138 138 MISSING (IN REF. 3).  
 FT CONFLICT 233 329 MISSING (IN REF. 3).  
 FT CONFLICT 421 421 MISSING (IN REF. 3).  
 SQ SEQUENCE 1447 AA; 158456 MW; 4A8612766ED0471F CRC64;

Query Match 6.5%; Score 136; DB 1; Length 1447;

Best Local Similarity 25.1%; Pred. No. 0.0072;

Matches 62; Conservative 34; Mismatches 101; Indels 50; Gaps 10;

QY 9 RDPGSGSGNEV-----IEGPNARVLKGSQARFNCTVSCQW--KLIMW 50

Db 215 RNPASRTGNEAEVRLSDPGLHRLYFLQPSNVVALEGKDAVECCVS-GYPPSPFTW 273

QY 51 ALSDMVTVSVRMEPIIINDRFTSQRDQGNFTSEMIHNVPDSGNIRC--SLQNSR 108

Db 274 LRGEVI-----QLRSKYSLLGG--SNLLISNVTDDSGMYTCVVIYKNN 318

QY 109 LHGSAYLTQVMGELFIPSVNLVVAENPECVTCLPSSHWTWLPDISW-ELGLLVSHSSYY 167

Db 319 ISASAEITLVLPFPFLNHPNLSYAYESMDIEFECTVSGKP-VPTVMNKNKGDVTPSDYF 377

QY 168 FVPEPSDQSAVILALTPQSNGLTCVATWTKSLKARKSATVNLTVIRCPQDTGGGINIP 227  
 Db 378 QIVGGSNLR-----ILGVKSDGFGYQCVAEAGNAQTSQAQLIVPRAIPSSS----- 426

QY 228 GVLSSLP 234  
 Db 427 -VLFSAP 432

RESULT 3  
 DCC\_MOUSE  
 ID\_DCC\_MOUSE STANDARD; PRT; 1447 AA.  
 AC P70211;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Tumor suppressor protein DCC precursor.  
 GN DCC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=96112625; PubMed=8570174;  
 RA Cooper H.M., Armes P., Britto J., Gad J., Wilks A.F.;  
 RT "Cloning of the mouse homologue of the deleted in colorectal cancer  
 gene (mDCC) and its expression in the developing mouse embryo.";  
 RL Oncogene 11:2243-2254 (1995).  
 RN [2]  
 RP REVISIONS.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RA Cooper H.M.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=A;  
 CC IsoId=P70211-1; Sequence=Displayed;  
 CC Note=Isoform B is produced by alternative initiation at Met-85  
 CC of isoform A;  
 CC Name=C;  
 CC IsoId=P70211-2; Sequence=VSP\_002501;  
 CC Event=Alternative initiation;  
 CC Comment=2 isoforms, A (shown here) and B, are produced by  
 CC alternative initiation at Met-1 and Met-85;  
 CC -!- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE  
 CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN  
 CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.  
 CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.  
 CC -!- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS  
 CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION  
 CC AND REMAIN AT THIS LEVEL IN THE ADULT.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.  
 CC -!- SIMILARITY: Contains 6 fibronectin type III domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X85788; CAA59786.1; -;  
 CC HSSP; PS6276; ITLK.  
 CC MGD; MGI:94869; DCC.  
 CC InterPro; IPR003961; FN\_III.

DR InterPro; IPR003962; FNIII subd.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00041; fn3; 6.  
 DR Pfam; PF00047; Ig; 4.  
 DR PRINTS; PR00014; FNTPETIII.  
 DR SMART; SM00060; FN3; 6.  
 DR SMART; SM00408; IGC2; 3.  
 DR PROSITE; PS0835; IG-LIKE; 4.  
 KM Glycoprotein; immunoglobulin domain; Transmembrane; Signal;  
 KM Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.  
 FT SIGNAL 1 25  
 FT CHAIN 26 1447  
 FT CHAIN 85 1447  
 FT INIT\_MET 85 85  
 FT DOMAIN 26 1097  
 FT TRANSMEM 1098 1122  
 FT DOMAIN 1123 1447  
 FT DOMAIN 36 135  
 FT DOMAIN 139 229  
 FT DOMAIN 234 326  
 FT DOMAIN 331 416  
 FT DOMAIN 426 522  
 FT DOMAIN 525 618  
 FT DOMAIN 619 716  
 FT DOMAIN 722 816  
 FT DOMAIN 840 940  
 FT DOMAIN 941 1042  
 FT DISULFID 61 117  
 FT DISULFID 161 212  
 FT DISULFID 261 310  
 FT DISULFID 352 400  
 FT CARBOHYD 60 60  
 FT CARBOHYD 94 94  
 FT CARBOHYD 299 299  
 FT CARBOHYD 318 318  
 FT CARBOHYD 478 478  
 FT CARBOHYD 628 628  
 FT CARBOHYD 702 702  
 FT VARSPLIC 819 838  
 FT  
 FT  
 SQ SEQUENCE 1447 AA; 158298 MW; 0D1F1097C22D5B9F CRC64;  
 Query Match 6.3%; Score 132; DB 1; Length 1447;  
 Best Local Similarity 25.5%; Pred. No. 0.015;  
 Matches 56; Conservative 31; Mismatches 91; Indels 42; Gaps 9;  
 QY 9 RPPPGSGGNEV-----TEGPNARVLKGSQARFNCTVSGW--KLIMW 50  
 Db 215 RNPASIRTGNEAEVRILSDPGLHRLQYFLQPSNVIAIEGKDAVLECCVS-GYPPPSFTW 273  
 QY 51 ALSDMVLVSVRPMEPIITNDRTSQRYDOGNFTSEMIHNVPSPSGNIRC--SLQNSR 108  
 Db 274 LRGEVI-----QURSKYSLGG--SNLLISNVYDDSGYTCVVTYKKN 318  
 QY 109 LHGSAYLTQVMGELFIPSNVLVAENPECVTCLPSHMTWLPDISW-ELGLLVSHSSYY 167  
 Db 319 ISASAEALTVLVPPWFLNHPNLSNLYAVESMDIECAVSGKP-VPTVNMKNKGDVIFSDYF 377  
 QY 168 FVPEPSDQSAVILALTPQSNGLTCTVATWTKSLKARKSA 207  
 Db 378 QIVGGSNLR-----ILGVKSDGFGYQCVAEAGNAQSA 413

RESULT 4  
 KML5\_HUMAN  
 ID KML5\_HUMAN STANDARD; PRT; 1914 AA.  
 AC Q15746; O95796; O95797; O95798; O95799; Q14844; Q16794; Q9UBG5;  
 AC Q9UIT9;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)













AC Q9P121; 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Neurotrophin precursor (hnt).  
GN NT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li G., Jin J., Tan X., Hu S., Yuan J., Qiang B.;  
RT "Cloning and identification of human neurotrophin full length cDNA";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Neural cell adhesion molecule.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF126426; AAF37591.1; --  
DR GO; GO:0008038; P:neuronal cell recognition; TAS.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003598; IG\_C2.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; ig; 3.  
DR SMART; SM00408; Igc2; 2.  
DR PROSITE; PS00835; IG\_LIKE; 3.  
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
KW Repeat; Signal.  
FT SIGNAL 1 31  
FT CHAIN 32 321  
FT PROPEP 322 344  
FT DOMAIN 39 126  
FT DOMAIN 136 218  
FT DOMAIN 222 309  
FT DISULFID 57 115  
FT DISULFID 157 201  
FT DISULFID 243 295  
FT CARBOHYD 44 44  
FT CARBOHYD 70 70  
FT CARBOHYD 152 152  
FT CARBOHYD 284 284  
FT CARBOHYD 292 292  
FT CARBOHYD 305 305  
FT CARBOHYD 321 321  
FT LIPID 321 321  
SQ SEQUENCE 344 AA; 37971 MW; DA4D12C295ABBE3A CRC64;  
  
Query Match 5.5%; Score 115; DB 1; Length 344;  
Best Local Similarity 23.7%; Pred. No. 0.057;  
Matches 58; Conservative 37; Mismatches 94; Indels 56; Gaps 13;  
  
QY 26 NARVLKGSQARFNCVTSQGWKLIMWLSVLS-----VRPMEPIITDRFTSQRYDQ 79  
DB 44 NTVVRQGESATLRCTIDNRTVRVAMINRSTIYAGNDKWLCDPRVVLN---TQTQY-- 98  
  
QY 80 GGNFTSEMTHNVERSDSNIRCSLQ-----NSRLHGSAYLTVMGELFIPSNLVVA 133  
DB 99 -----SIBIQNVVDYDEGFTVCSTQTDNHPKTSRVH-----LIVQVSPKIVEISSDISIN 148  
  
QY 134 ENPECEVTCL---PSHTWTLDPDISWELGLLVSHSSYFVPEPSDLSQASVIALTPQSN 189  
DB 149 EGNISLTCTATGRPE-----FTVWR---HISPKAVGFVSEDEYLE---IQITREQS 196

QY 190 GLTCTVATWTKSLKARKSATVNLTVIRCP-----ODTGGGINIPGVL-----SSLPSLGFSL 240  
DB 197 GUYECAS-NDVAAPVRRVKVTVNYPYISEAKGTGVPVGKGLTQCEASAVPSAEFQ- 254  
QY 241 PTWGX 245  
DB 255 --WYK 257  
  
RESULT 11  
NTRI\_MOUSE  
ID NTRI\_MOUSE STANDARD; PRT; 344 AA.  
AC Q99F07;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Neurotrophin precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=ICR; TISSUE=Brain;  
RA Kim T.H., Choi S.C., Kim J., Jeon J.W., Kim K.D., Lee S.H.;  
RT "Cloning and expression of mouse neurotrophin gene in the developing  
RT nervous system";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN=C57BL/6; TISSUE=Eye;  
RA MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Neural cell adhesion molecule.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF282980; AAK00276.1; --  
DR EMBL; BC023307; AAH23307.1; --  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR003598; IG\_C2.  
DR InterPro; IPR003006; IG\_MHC.







```

Genomics 41:414-421(1997).
CC -!- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q92859-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q92859-2; Sequence=VSP_002593;
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED AND ALSO IN CANCER CELL
CC LINES.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 6 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U61262; AAB17263.1; -;
DR EMBL; U72391; AAC51287.1; -;
DR HSP; P02751; 1TTF.
DR Genew; HGNC:7754; NEOL.
DR MIM; 601907; -;
DR GO; GO:0003687; C:integral to plasma membrane; TAS.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; ig; 4.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS00835; IG_LIKE; 4.
KW Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;
KW Glycoprotein; Alternative splicing.
FT SIGNAL 1 33
FT CHAIN 34 1461
FT DOMAIN 34 1105
FT DOMAIN 1106 1126
FT TRANSMEM 1127 1461
FT DOMAIN 52 141
FT DOMAIN 152 238
FT DOMAIN 243 336
FT DOMAIN 341 426
FT DOMAIN 436 533
FT DOMAIN 536 629
FT DOMAIN 630 729
FT DOMAIN 735 829
FT DOMAIN 850 950
FT DOMAIN 951 1052
FT DOMAIN 1118 1121
FT DISULFID 74 129
FT DISULFID 173 221
FT DISULFID 270 320
FT DISULFID 362 410
FT CARBOHYD 73 73
FT CARBOHYD 210 210
FT CARBOHYD 326 326
FT CARBOHYD 470 470
FT CARBOHYD 489 489
FT CARBOHYD 639 639

```

Genomics 41:414-421(1997).  
-!- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE  
TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR  
DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION  
MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.  
-!- SUBCELLULAR LOCATION: Type I membrane protein.  
-!- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Comment=Additional isoforms seem to exist;  
Name=1;  
IsoId=Q92859-1; Sequence=Displayed;  
Name=2;  
IsoId=Q92859-2; Sequence=VSP\_002593;  
-!- TISSUE SPECIFICITY: WIDELY EXPRESSED AND ALSO IN CANCER CELL  
LINES.  
-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC  
SUBFAMILY.  
-!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.  
-!- SIMILARITY: Contains 6 fibronectin type III domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; U61262; AAB17263.1; -;  
EMBL; U72391; AAC51287.1; -;  
HSP; P02751; 1TTF.  
Genew; HGNC:7754; NEOL.  
MIM; 601907; -;  
GO; GO:0003687; C:integral to plasma membrane; TAS.  
InterPro; IPR003961; FN\_III.  
InterPro; IPR003962; FNIII subd.  
InterPro; IPR007110; IG-like.  
InterPro; IPR003598; IG\_c2.  
InterPro; IPR003006; IG\_MHC.  
Pfam; PF00041; fn3; 6.  
Pfam; PF00047; ig; 4.  
PRINTS; PR00014; FNTYPEIII.  
SMART; SM00060; FN3; 6.  
SMART; SM00408; IGC2; 3.  
PROSITE; PS00835; IG\_LIKE; 4.  
KW Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;  
KW Glycoprotein; Alternative splicing.  
FT SIGNAL 1 33  
FT CHAIN 34 1461  
FT DOMAIN 34 1105  
FT DOMAIN 1106 1126  
FT TRANSMEM 1127 1461  
FT DOMAIN 52 141  
FT DOMAIN 152 238  
FT DOMAIN 243 336  
FT DOMAIN 341 426  
FT DOMAIN 436 533  
FT DOMAIN 536 629  
FT DOMAIN 630 729  
FT DOMAIN 735 829  
FT DOMAIN 850 950  
FT DOMAIN 951 1052  
FT DOMAIN 1118 1121  
FT DISULFID 74 129  
FT DISULFID 173 221  
FT DISULFID 270 320  
FT DISULFID 362 410  
FT CARBOHYD 73 73  
FT CARBOHYD 210 210  
FT CARBOHYD 326 326  
FT CARBOHYD 470 470  
FT CARBOHYD 489 489  
FT CARBOHYD 639 639

FT CARBOHYD 715 715 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 909 909 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 1248 1300 Missing (in isoform 2).  
FT FTId=VSP\_002593.  
FT CONFLICT 168 168 G -> N (IN REF. 2).  
SQ SEQUENCE 1461 AA; 159958 MW; 7AAE897E69635A21 CRC64;  
Query Match 5.4%; Score 113; DB 1; Length 1461;  
Best Local Similarity 23.1%; Pred. No. 0.5;  
Matches 65; Conservative 34; Mismatches 110; Indels 72; Gaps 12;  
QY 18 NEVIEGPQNAVRL-----KGSQARFNCVTSQGWKLIMWALSDMVLSVRPMEPIITN---- 69  
DB 196 DRVIKPLSGMLVISNATEGDGLYRCVVESSGGPP---KYSDEVELKVLDPPEVSDLVFL 252  
QY 70 -----DFTSQRYD--QGNGTSEMIH 90  
DB 253 KQPSPLVRVIGQDVLPVPCVAGSLPTTIKMKNEEALDTESSRLVLAGGS----LEIS 308  
QY 91 NVEPDSGNIRCSQNSR--LHGSAYLTQVMGEFLFIPSVNLVVAENE-----PCEVTCLP 144  
DB 309 DVTEDDAGTYFCIADNGNETIEAQAELTVQAQPEFLKQPTNIYAHESMDIVFECEVTGKP 368  
QY 145 SHWTWLPDISW-ELGLLVSHSSYFVFPESDLSQSAVSIILATPQSNGLTLCVATWKSLSKA 203  
DB 369 T-----PTVKWKGMDVIPSDYFKIVKHNLQ-----VLGLVKSDRGFYQCI AENDVGN 419  
QY 204 RKSATVNLTVIRCPQDTGGGINP-GVLSLPSLGFSLPTW 243  
DB 420 QAGAQL-IILEHAPATTGPLSPAPRDVVASLVSTRFIKLTW 459  
RESULT 15  
OPCM HUMAN  
ID OPCM HUMAN STANDARD; PRT; 345 AA.  
AC Q1492;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Opioid binding protein/cell adhesion molecule precursor (OBCAM)  
DE (Opioid-binding cell adhesion molecule) (OPCML).  
GN OPCML OR OBCAM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]\_SEQUENCE FROM N.A.  
RP SEQUENCE FROM N.A.  
RC TISSUE=Occipital cortex;  
RX MEDLINE=95237612; PubMed=7721093;  
RA Shark K.B., Lee N.M.;  
RT "Cloning, sequencing and localization to chromosome 11 of a cDNA  
RL encoding a human opioid-binding cell adhesion molecule (OBCAM).";  
Gene 155:213-217(1995)  
CC -!- FUNCTION: Binds opioids in the presence of acidic lipids; probably  
involved in cell contact.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By  
similarity).  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L34774; AAA36387.1; -;  
DR PIR; JC4025; JC4025.  
DR Genew; HGNC:8143; OPCML.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:31:20 ; Search time 12.3761 Seconds  
(without alignments)  
1319.641 Million cell updates/sec

Title: US-09-729-264-4  
Perfect score: 2088  
Sequence: 1 MVAGAMENRPPGSGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pdp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pdp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pdp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pdp.\*  
5: /cgn2\_6/ptodata/2/iaa/6CTUS\_COMB.pdp.\*  
6: /cgn2\_6/ptodata/2/iaa/6CTUS\_COMB.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	145.5	7.0	404	4	US-09-638-649-3
2	136	6.5	1447	3	US-09-041-886-25
3	136	6.5	1447	5	PCT-US94-05277-2
4	132.5	6.3	869	1	US-08-374-834-16
5	132.5	6.3	869	2	US-08-644-271-29
6	132.5	6.3	869	4	US-09-077-955-33
7	128.5	6.2	1345	2	US-08-977-767-3
8	123.5	5.9	340	4	US-09-651-200-2
9	123.5	5.9	441	4	US-09-651-200-4
10	122.5	5.9	534	4	US-09-651-200-6
11	122.5	5.9	534	4	US-09-651-200-24
12	121.5	5.8	332	4	US-09-062-365-1
13	120	5.7	1395	3	US-08-540-245A-15
14	118	5.7	868	1	US-08-374-834-1
15	118	5.7	868	2	US-08-644-271-1
16	118	5.7	868	4	US-09-077-955-1
17	116.5	5.6	318	2	US-08-633-148-4
18	116.5	5.6	340	2	US-08-633-148-2
19	115.5	5.5	478	5	PCT-US95-08493-15
20	115.5	5.5	860	5	PCT-US95-08493-19
21	115.5	5.5	868	5	PCT-US95-08493-21
22	115.5	5.5	946	5	PCT-US95-08493-13
23	113	5.4	325	4	US-09-651-200-20
24	112.5	5.4	416	4	US-09-638-649-1
25	112	5.4	689	4	US-08-499-364-1
26	110	5.3	352	4	US-09-996-243-505
27	109	5.2	421	2	US-08-659-984A-1

ALIGNMENTS

RESULT 1

US-09-638-649-3  
; Sequence 3, Application US/09638649  
; Patent No. 6563015  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Yan, Shi Du  
; TITLE OF INVENTION: TRANSGENIC MICE OVER-EXPRESSING RECEPTOR FOR ADVANCED GLYCATION ENDPRODUCT (RAGE) AND MUTANT APP IN BRAIN AND TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 0575/62175  
; CURRENT APPLICATION NUMBER: US/09/638,649  
; CURRENT FILING DATE: 2000-08-14  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Human  
US-09-638-649-3

Query Match 7.0%; Score 145.5; DB 4; Length 404;

Best Local Similarity 23.5%; Pred. No. 5.5e-05;

Matches 77; Conservative 34; Mismatches 94; Indels 123; Gaps 15;

Qy	64	EPIITNDRFTS-----QRYDQGGNFT--SEMIHNVFSDSGNIR-----CSLQNSRLHGS	112
Db	162	KPLVNEKGVSVKQTRRHPETGLTLQSELM---VTTPARGGDRPTFSCSFSGLPGRHR	218
Qy	113	AYLTVMQVNGELFIP-----SVNLVVAENP-----CEVTCLPSSHWTWLPDIS	154
Db	219	ALRTAPIQPRWEVPVLEEVQLVV---EPEGGAVAGGTVTLTCEVPAQPS-----PQIH	270
Qy	155	WEGLLVSHSSYFVPFVPEFSDLSQAVSILALTPQSNGLTCTVATWKLKARKSATVNLTVI	214
Db	271	WMKD-----GVPLPLPSPVLIIPETGPQDQGYSCVATHSSHPQBSRAVSIIL	321
Qy	215	RCQDQDGGGINIPGVLSLSLPSLOFSLPTWGVKGLAGTMLLT-----PTCTLTIRCC	267
Db	322	E-FGEEG-----PTAGSVGSGGLGTALALGILGGLGTAALLIGVI	361
Qy	268	CCRRRCGCGCCGCCCCCKRKGFRIOQKKSEKXT--NKETETSGNENSGYNDEQ	325
Db	362	LMQRR-----QRRGEERKAPENQBEERAEALN-----	389
Qy	326	KTTDTASLPKPKSCSSDPEQNRSSCGPP	353
Db	390	-----QSEPEAGESSTGGP	404

## RESULT 2

US-09-041-886-25  
; Sequence 25: Application US/09041886  
; Patent No. 6235872  
; GENERAL INFORMATION:  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Rabizadeh, Sharro  
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence  
; TITLE OF INVENTION: Polypeptides and Methods of Use  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/041.886  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2626  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US94-05277-2

Query Match 6.5%; Score 136; DB 3; Length 1447;

Best Local Similarity 25.1%; Pred. No. 0.0023;

Matches 62; Conservative 34; Mismatches 101; Indels 50; Gaps 10;

QY	9	RDPFGSGNEV-----IEGPQNAVRLKGSQARFNCVTQSGM--KLIMW	50
Db	215	RNPASRTGNEAEVRILSDPGLHRLQFLQRPNSNVAIEGKDAVLECCVS-GYPPPSFTW	273
QY	51	ALSDMVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVPSDSGNIRC--SLNSR	108
Db	274	LRGEEVI-----QLRSKYSLLGG--SNLLISNVTDSDSGMTVCVVYTKNEN	318
QY	109	LHGSAYLTVQVMGELFIPSVNLVVAENPEVCITCLPSHWTWLPDISW-ELGLLVSHSYY	167
Db	319	ISASAEITLVPPFWFLNHPNSNLAYESMDIEFFECTVSGKP-VPTVNMKNGDVVIPSDF	377
QY	168	FVPEPDLQSAVSIILALTPOSNGTLTCVATWKSILKARKSATVNLTVIRCQDGTGGGINP	227
Db	378	QIVGGSNLR---ILGVKSDGFIQCVAEANEAGNAQTSALQILVPKPAIPSSS-----	426
QY	228	GVLSLSP 234	
Db	427	-VLPSAP 432	

## RESULT 3

PCT-US94-05277-2

; Sequence 2: Application PC/TUS9405277

; GENERAL INFORMATION:

; APPLICANT: Bruskin, Arthur

; APPLICANT: Jarosz, David E.

; APPLICANT: Johnson, Karen  
; APPLICANT: Kinzler, Kenneth W.  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Zablach, James R.  
; TITLE OF INVENTION: Antibodies Specific for DCC Gene Product  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie & Beckett  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05277  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.42709  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202.508.9100  
; TELEFAX: 202.508.9299  
; TELEX: 197430 BEMB UT  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US94-05277-2

Query Match 6.5%; Score 136; DB 5; Length 1447;

Best Local Similarity 25.1%; Pred. No. 0.0023;

Matches 62; Conservative 34; Mismatches 101; Indels 50; Gaps 10;

QY	9	RDPFGSGNEV-----IEGPQNAVRLKGSQARFNCVTQSGM--KLIMW	50
Db	215	RNPASRTGNEAEVRILSDPGLHRLQFLQRPNSNVAIEGKDAVLECCVS-GYPPPSFTW	273
QY	51	ALSDMVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVPSDSGNIRC--SLNSR	108
Db	274	LRGEEVI-----QLRSKYSLLGG--SNLLISNVTDSDSGMTVCVVYTKNEN	318
QY	109	LHGSAYLTVQVMGELFIPSVNLVVAENPEVCITCLPSHWTWLPDISW-ELGLLVSHSYY	167
Db	319	ISASAEITLVPPFWFLNHPNSNLAYESMDIEFFECTVSGKP-VPTVNMKNGDVVIPSDF	377
QY	168	FVPEPDLQSAVSIILALTPOSNGTLTCVATWKSILKARKSATVNLTVIRCQDGTGGGINP	227
Db	378	QIVGGSNLR---ILGVKSDGFIQCVAEANEAGNAQTSALQILVPKPAIPSSS-----	426
QY	228	GVLSLSP 234	
Db	427	-VLPSAP 432	

## RESULT 4

US-08-374-834-16

; Sequence 16: Application US/08374834

; Patent No. 5656473

; GENERAL INFORMATION:

; APPLICANT: Valenzuela, et al.

; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Regeneron Pharmaceuticals, Inc.



Qy	369	QASENLA	375
		: : :	
Db	483	SSSFVS	489

## RESULT 6

US-09-077-955-33  
; Sequence 33, Application US/09077955A  
; Patent No. 6413740

## ; GENERAL INFORMATION:

```

; APPLICANT: Valenzuela et al., David M.
;
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
;
; FILE REFERENCE: REG195-B-PCT-US
;
; CURRENT APPLICATION NUMBER: US/09/077,955A
;
; CURRENT FILING DATE: 1998-09-10
;
; EARLIER APPLICATION NUMBER: PCT/US96/20696
;
; EARLIER FILING DATE: 1996-12-13
;
; EARLIER APPLICATION NUMBER: 08/644,271
;
; EARLIER FILING DATE: 1996-05-10
;
; EARLIER APPLICATION NUMBER: 60/008,657
;
; EARLIER FILING DATE: 1995-12-15
;
; NUMBER OF SEQ ID NOS: 36
;
; SOFTWARE: Patent Ver. 2.0

```

Query Match 6.3%; Score 132.5; DB 4; Length 869;  
Best Local Similarity 20.8%; Pred. No. 0.0023;  
Matches 89; Conservative 55; Mismatches 178; Indels 105

QY 63 MEPIITNDRFTSQRYDQGNFTSEMIHNVEPDSGNIRCSLQNSRLHGSAY-LTVQWVG 121

Db 159 -SDIPENSPAVLE-----SGSPITHNVKEDRSGCVNATVAVSVC-----GMSVGNFTFTY 200

**OY**

122 ELFIPSVNLVAENEP-----CEVTCLPSTHWTLDPISW-ELGILLVSHSSYYFFPEP 172  
| : | : | : | : | : | : | : | : |  
**Dd** 209 EVFARIILRAPESHNVTFGSGFTVLHCTATGP----VPTTWIENGNAVSGSIOBSVKD 263

QY            173 S L Q S A V S I L A L T P Q N G F L T C V A T --- W K S L K A R K S A T V N L T V I R C P Q D T G G G ----- 223  
               : : : : | : : : | : : : | : : : | : : : |  
Db            264 R V I D S R L Q F I T X P --- G L Y T C I A T N H G E K F S T A K A A A T I S I A E W S K P Q D N K G I C A Q Y 320

QY 224 ----INFGVLSSELSFGTSL-----PWGRVGGGLAGHMLLTPTCTLTATRCQ 267

Db 321 RGEVCNAVIAKDALVFLNTSYADPEEAQELLVHTAWNEL-----KVSVPCRPAAEAL 373

Db 374 LCNHIFQCSPGVPTPIPICREYCLAVKELFCAKE---WLVMEEKTHRGLYRSEMILL 429

Db 430 SVPECSKLPSMHWDTACARLP-----HLDYNKENLKTFFP--MTSSKPSVDIPNLPS 482

Qv 369 CASNTA 375

RESULT 7

US-08-977-767-3  
; Sequence 3, Application US/08977767  
; Patent No. 5972684  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga

APPLICANT: Yue, Henry  
APPLICANT: Greenwald, Sara  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: CARBONIC ANHYDRASE VIII  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/977,767  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

Query Match 6.2%; Score 128.5; DB 2; Length 1345;  
Best Local Similarity 33.0%; Pred. No. 0.0098;  
Matches 37; Conservative 1; Mismatches 41; Indels 33; Gaps 5.

Qy 246 VGLGLAGTMLLT-PTCLTLIRCCCCRRRCGCCGCCRC-----CFCC 286

Db 456 CGTGAAGAGGGTCCTCTATGACCCCTCTGCCCCCTCTGAGACTCAGCACC 507

## RESULT 8

```

US-09-051-200-2
; Sequence 2, Application US/09651200
; Patent No. 6429303
; GENERAL INFORMATION:
; APPLICANT: Green et al
; TITLE OF INVENTION: Polynucleotides Encoded
; TITLE OF INVENTION: Lymphocyte Activation
; TITLE OF INVENTION: Polypeptides Encoded
; FILE REFERENCE: 15966-562 (CURA-62)
; CURRENT APPLICATION NUMBER: US/09/651,200
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152383
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/172909
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/183578
; PRIOR FILING DATE: 1999-12-21

```

```
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-651-200-2

Query Match
Best Local Similarity 5.9%; Score 123.5; DB 4; Length 340;
Matches 78; Conservative 50; Mismatches 130; Indels 109; Gaps 17;

QY 11 PPGSGGNEVIEGPONARV-LKGSQAREFNCTVS--CGWKL-----IMWALSDMWVLSVRPM 63
Db 45 PORSPTGAVEVQVPEDPVVALVGTDLHCFSFEPGFSLTQLNLWQDTDTKQLV-----201
QY 64 EPIITNDRFTSQRYDQGNF-----TSEMIHNVPEPSDSGNIRCSLQNSRL 109
Db 202 -----HSFTEGR-DQGSAYANRTALFPDLLAQNASLRQVRVVADEGSFTCFV-SIRD 253
QY 110 HGSAYLTVQVMGELFIPSNLV-----VAENPECEVTCLPSHWTWLP--DISWELGL---L 160
Db 254 FGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDGGVPL 311
QY 161 VSHSSYFVPEPSDLSQSAVSIILALTPQSNGLTLCVATWKS LKARKSATVNLTVIRCP---217
Db 312 TGNVTTSQMANEGGLFDVHSLRVVLGANGTYSC-----LVRNPVLQ 353
QY 218 QDTGGGINPGVLSLPSLGSFSLPTWKGVLGLAGTMLLTPTCTLTIRCCCCRRCCGCN 277
Db 354 QDAHGSVTTTGQPMTPPEAL-----WVTVGLSVCLLALLV-----389
QY 278 CCRCFCFCRRKRGFRQFKKSEKTKETETESGNENSGYNSDEQKTTDTASLPKKS 337
Db 390 ---ALAFVCWRK-----IKQSCSEENAGAEDQDG-----EGGSKTALQPLKH 429
QY 338 CESSDPE 344
Db 430 SDSKEDD 436

RESULT 10
US-09-651-200-6
; Sequence 6, Application US/09651200
; Patent No. 6429303
; GENERAL INFORMATION:
; APPLICANT: Green et al
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-562 (CURA-62)
; CURRENT APPLICATION NUMBER: US/09/651,200
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152383
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/172909
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/183578
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-651-200-6

Query Match
Best Local Similarity 5.9%; Score 122.5; DB 4; Length 534;
Matches 78; Conservative 49; Mismatches 131; Indels 109; Gaps 17;

QY 11 PPGSGGNEVIEGPONARV-LKGSQAREFNCTVS--CGWKL-----IMWALSDMWVLSVRPM 63
Db 239 PORSPTGAVEVQVPEDPVVALVGTDLHCFSFEPGFSLTQLNLWQDTDTKQLV-----294
QY 64 EPIITNDRFTSQRYDQGNF-----TSEMIHNVPEPSDSGNIRCSLQNSRL 109
Db 295 -----HSFTEGR-DQGSAYANRTALFPDLLAQNASLRQVRVVADEGSFTCFV-SIRD 346
QY 110 HGSAYLTVQVMGELFIPSNLV-----VAENPECEVTCLPSHWTWLP--DISWELGL---L 160
Db 347 FGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDGGVPL 404
QY 161 VSHSSYFVPEPSDLSQSAVSIILALTPQSNGLTLCVATWKS LKARKSATVNLTVIRCP---217
Db 405 TGNVTTSQMANEGGLFDVHSLRVVLGANGTYSC-----LVRNPVLQ 446
```

QY 218 QDTGGGINIPGVLSLPSLGSFSLPTWKGVLGLAGTMLLTPTCTLTIRCCRRCCGCGN 277  
Db 447 QDAHGSVTTIGQPMTFPEAL-----WVTGLSVCLIALLV----- 482  
QY 278 CCRCFCFCCRRKRGFRIQKQSEKTKNETETESGNENSGVNSDEQKTTDTASLPKKS 337  
Db 483 ---ALAFVCRK-----IKQSCSEENAGAEADQD-----EGGSKTALQPLKH 522  
QY 338 CESSDPE 344  
Db 523 SDSKEDD 529  
RESULT 11  
US-09-651-200-24  
; Sequence 24, Application US/09651200  
; Patent No. 6429303  
; GENERAL INFORMATION:  
; APPLICANT: Green et al  
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B  
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 15966-562 (CURA-62)  
; CURRENT APPLICATION NUMBER: US/09/651,200  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/152383  
; PRIOR FILING DATE: 1999-09-03  
; PRIOR APPLICATION NUMBER: 60/172909  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/183578  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Sequence  
; OTHER INFORMATION: mz5020.protein from Figure 4.  
US-09-651-200-24

Query Match 5.9%; Score 122.5; DB 4; Length 534;  
Best Local Similarity 21.3%; Pred. No. 0.0095;  
Matches 78; Conservative 49; Mismatches 131; Indels 109; Gaps 17;  
QY 11 PFGSGSGNEVIEGPQNAV-LKGSQARFNCTVS--QGWKL-----IMWLSDMVLSVRPM 63  
Db 239 PQRSPTGAVEVQVPEPVVALVGTDLRCSFSPFGLAQNLILWLTDTKQLV----- 294  
QY 64 EPIITNDRTSQRYDGGNF-----TSEMIHNVEPSDSGNIRCSLQNSRL 109  
Db 295 -----HSTEGE-DQGSAYANTALFPDLAQGNASRLQRVRVADEGSFTCFV-SIRD 346  
QY 110 HGSAYLTVOVMGELFPPSNLV-----VAENPECVTLCPSHWTWLP--DISWELGL--L 160  
Db 347 FGSAAVSLQVAPYKPSFTLEPNKDLRPGDTVTIIC--SSVGYPEAEVFWDDGGVFL 404  
QY 161 VHSYFYFPEPSDLSQAVSIILALTPQSNGLTLCVATWKSLSKARKSATVNLTVIRCP-- 217  
Db 405 TGNVTTSQMANEQGLFVHSLRVLGLANGTYS-----LVRNPVLQ 446  
QY 218 QDTGGGINIPGVLSLPSLGSFSLPTWKGVLGLAGTMLLTPTCTLTIRCCRRCCGCGN 277  
Db 447 QDAHGSVTTIGQPMTFPEAL-----WVTGLSVCLIALLV----- 482  
QY 278 CCRCFCFCCRRKRGFRIQKQSEKTKNETETESGNENSGVNSDEQKTTDTASLPKKS 337  
Db 483 ---ALAFVCRK-----IKQSCSEENAGAEADQD-----EGGSKTALQPLKH 522  
QY 338 CESSDPE 344  
Db 523 SDSKEDD 529

Db 523 SDSKEDD 529  
RESULT 12  
US-09-062-365-1  
; Sequence 1, Application US/09062365  
; Patent No. 6465422  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Stern, David  
; TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A  
; TITLE OF INVENTION: SUBJECT  
; FILE REFERENCE: 55424  
; CURRENT APPLICATION NUMBER: US/09/062,365  
; CURRENT FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Human  
US-09-062-365-1  
Query Match 5.8%; Score 121.5; DB 4; Length 332;  
Best Local Similarity 25.7%; Pred. No. 0.0061;  
Matches 57; Conservative 27; Mismatches 69; Indels 69; Gaps 11;  
QY 64 EPIITNDRTS-----QRYDQGGNFT--SEMIHNVEPSDSGNIR-----CSLQNSRLHGS 112  
Db 140 KFLVFNKGVSKVKEQTRRHPTGLFTLQSELM---VTPARGGDPRTFTSCSFGLPRHR 196  
QY 113 AYLTVQVMGELFIP-----SVNLVVAENEP-----CEVTCLPSTWTLPLDIS 154  
Db 197 ALRTAPIQPRVWPEVPLEEVQLVV---EPEGGAVAPGGVTTLTCEVPAQPS-----PQIH 248  
QY 155 WELGLLVSHSSYFYFPEPSDLSQAVSIILALTPQSNGLTLCVATWKSLSKARKSATVNLTVI 214  
Db 249 WMKD-----GVPLPLPPSVLILPEIGPDQGTYSVATHSSHGQESRAVSISII 299  
QY 215 RCPQDGTGGGINIPGVLSLPSLGSFSLPTWKGVLGLAGTMLL 256  
Db 300 E-PGEEG-----PTAGSVGGSLGTIAL 321

RESULT 13  
US-09-540-245A-15  
; Sequence 15, Application US/09540245A  
; Patent No. 6270984  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; APPLICANT: Tessier-Lavigne, Marc  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,245A  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1997-11-14  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 1395  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-540-245A-15  
Query Match 5.7%; Score 120; DB 3; Length 1395;  
Best Local Similarity 26.2%; Pred. No. 0.06;  
Matches 66; Conservative 29; Mismatches 101; Indels 56; Gaps 15;

QY 24 PQNARVLKGSQARPNCVTSG--WKLMW-----ALSDMVVLVSRMEPIITNDRFTSQR 76  
 DB 161 PKDTRVAKGATALJECGPPKIGIPEPTLIWIKDGVPLDDKAMFGASSRV-----R 211  
 QY 77 YDQGNFTSMIHNVPESDGNIRCSLON---SRLHGSAYLTVQVMGELFIPSVNLVVA 133  
 DB 212 IVDDGN---LLISNVEPIDEGNYKCLAQNLVGTRESSYAKLIVQKPYMKPKQOVML 267  
 QY 134 ENBP-----CEVTLCPSHWTLWDISW---ELGLLVSHSSYFFVPEPSDLSQSAVSIILALTP 186  
 DB 268 YGQTATFHCSVGGDPP-----PKVLWKKEGNIPIVSRARILHD-----EKSLEISNITP 316  
 QY 187 QSNGLTLTCAVA--TWKSLKARKSATV---NLTWIRCPQDTGGGINIPGVLSLSPSLGFSL 240  
 DB 317 TDEGTVCAHNNVGGISARASLIVHAPPNFT--KRPNSKKVGLN--GVV-QLPCMASGN 371  
 QY 241 PT-----WGKVL 248  
 DB 372 PPSVFTREGV 383

RESULT 14  
 US-08-374-834-1  
 ; Sequence 1, Application US/08374834  
 ; Patent No. 5656473  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Valenzuela, et al.  
 ; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
 ; STREET: 777 Old Saw Mill River Road  
 ; CITY: Tarrytown  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10591

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/374,834  
 ; FILING DATE: 19-JAN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/095,658  
 ; FILING DATE: 21-JUL-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Cobert, Robert J.  
 ; REFERENCE/DOCKET NUMBER: REG 190A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (914) 345-7400  
 ; TELEFAX: (914) 345-7721  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 868 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 US-08-374-834-1

Query Match 5.7%; Score 118; DB 1; Length 868;  
 Best Local Similarity 21.0%; Pred. No. 0.047;  
 Matches 83; Conservative 42; Mismatches 151; Indels 120; Gaps 19;  
 QY 4 GARNRDPGSGNEVIEGPONARVLKGSQARENCTVSGWK-LIMWALSDMVVLVSRP 62  
 DB 113 GALQVKMKP-----KIRPPINVKIIEGLKAVLPCTTNGNPKPSVSWTKGDSALRE--- 163  
 QY 63 MEPIITNDRFTSQRYDQGNFTSEMIHNVPESDGNIRCSLQNSRLHGSAY-LTVQVWG 121

DB 164 -----NSRIATVLE-----SGSLRIHNQKEDAGQYRCVAKNSL--GTAYSXLVKLEV 208  
 QY 122 ELFIPSVNLVVAENP-----CEVTLCPSHWTLWDISW-ELGLLVSHSSYFFVPEP 172  
 DB 209 EVFARILRAPESINVTGFSVTLRCTAIGMP-----VPTISWIENGNAVSSGSIQENVKD 263  
 QY 173 SDLSQSAVSIILALTPQSNGLTLTCAVA---WKSILKARKSATVNLTVIR-----CPQD 219  
 DB 264 RVIDSRQLFITKP---GLVTCIATNKHGKFKSTAKAAATVSIAEWSKSOKESKGYCAQY 320  
 QY 220 TGGGINIPGVLSLSPSLGFSLP-----TWGKVLGLAGTLMLLTPTCTLTTRCC 267  
 DB 321 RGEVDAVLVKOSLVFNTSYDPPEAEQELLIIHTAMNEL-----KAVSPLCRPAAEAL 373  
 QY 268 CRRRCRCGCN-----CCRCRC-----FCRRKRGRFRIQFQKKSEKKTNKETETE 312  
 DB 374 LCNHLFQECSPGVLPMPICREYCLAVKELFCA-----KEWLAMEGKTH 418  
 QY 313 SGNENSGYNSDEQKTTDTASLPKSCES-----SDP 343  
 DB 419 RGLYRSGMH-----FLPVPECSKLPSMHQDP 444

RESULT 15  
 US-08-644-271-1  
 ; Sequence 1, Application US/08644271  
 ; Patent No. 5814478  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Valenzuela, et al.  
 ; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS  
 ; TITLE OF INVENTION: AND LIGANDS  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
 ; STREET: 777 Old Saw Mill Road  
 ; CITY: Tarrytown  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10591  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/644,271  
 ; FILING DATE: 10-MAY-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: USSN 60/008,657  
 ; FILING DATE: 15-DEC-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Cobert, Robert J.  
 ; REFERENCE/DOCKET NUMBER: REG 195A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 914-345-7400  
 ; TELEFAX: 914-345-7721  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 868 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 US-08-644-271-1

Query Match 5.7%; Score 118; DB 2; Length 868;  
 Best Local Similarity 21.0%; Pred. No. 0.047;  
 Matches 83; Conservative 42; Mismatches 151; Indels 120; Gaps 19;

```
QY 4 GAMENRDPFGSGGNEVIEGPONARVLKGSQARFNCTVSGQWK-LIMWALSMDVVLVSRP 62
Db 113 GALQVKMP-----KITRPINVKIIEGLKAVLPCTTNGNPKPSVSWIKGDSALRE--- 163
QY 63 MEPIITNDRFTSQRYDQGNFTSEMIHNVPSDSGNIRCSLQNSRLHGSAY-LTVQVMG 121
Db 164 -----NSRIAYLE-----SGSLRIHNQKEDAGQYRCVAKNSL--GTAYSKLKLEV 208
QY 122 ELFIPSVNLVWAENRP-----CEVTCPLPSHTWMLPDISW-ELGLLVSHSYFYVPEP 172
Db 209 EVFARILRAPESHNVTFSGSVTLRCTAIGMP-----VFTISWENGNAVSSGSIQENVKD 263
QY 173 SDLQSAVSILALTPQSNGLTLCVAT----WKS LKARKSATVNLTVIR-----CPQD 219
Db 264 RVIDSRLQLFITKP--GLYTCIATNKHGEKFPSTAKAAATVSIAEWSKSKESKGCAQY 320
QY 220 TGGGINIFGVLSLPSLGSFLP-----TWGKVGLGLAGTMLLTPTCTLTIRCC 267
Db 321 RGEVCDAVLVKDSLVEFTNTSYDPDEAOELLIIHTAWNEL-----KAVSPLCRPAEAL 373
QY 268 CRRRCRCGCN-----CCRCRC-----FCCRRKRGRFIOFKKSEKTKETETE 312
Db 374 LCNHLFQECSPGVLPMPICREYCLAVKELFCA-----KEWLANEGKTH 418
QY 313 SGNENSGYNSDEQKTTDTASLPKSCS-----SDP 343
Db 419 RGLYRSGMH-----FLVPPEC SKLPSMHQDP 444
```

Search completed: November 13, 2003, 03:35:12  
Job time : 13.3761 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:34:36 ; Search time 26.7591 Seconds  
(without alignments)  
2477.484 Million cell updates/sec

Title: US-09-729-264-4

Perfect score: 2088

Sequence: 1 MVAGAMENRDPGSGSGNEV.....HPQAFNLASPEKYSNTTVV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	140	6.7	405	8	US-08-755-235-4
2	138.5	6.6	2473	15	US-10-184-644-559
3	138.5	6.6	2473	15	US-10-184-634-559
4	138	6.6	1477	12	US-10-274-583-20
5	138	6.6	1496	12	US-10-021-660-125
6	132.5	6.3	869	14	US-10-016-283-33
7	132	6.3	633	12	US-10-180-410-26
8	131.5	6.3	869	10	US-09-817-487A-2
9	128	6.1	592	12	US-10-180-410-2
10	128	6.1	594	12	US-10-180-410-12
11	128	6.1	708	12	US-10-180-410-8
12	128	6.1	708	12	US-10-199-672-584
13	128	6.1	708	12	US-10-187-749-584
14	128	6.1	708	12	US-10-194-457-584
15	128	6.1	708	12	US-10-184-642-584

16	128	6.1	708	12	US-10-196-747-584	Sequence 584, App
17	128	6.1	708	12	US-10-173-689-584	Sequence 584, App
18	128	6.1	708	12	US-10-173-690-584	Sequence 584, App
19	128	6.1	708	12	US-10-173-691-584	Sequence 584, App
20	128	6.1	708	12	US-10-173-692-584	Sequence 584, App
21	128	6.1	708	12	US-10-173-694-584	Sequence 584, App
22	128	6.1	708	12	US-10-173-698-584	Sequence 584, App
23	128	6.1	708	12	US-10-173-699-584	Sequence 584, App
24	128	6.1	708	12	US-10-173-707-584	Sequence 584, App
25	128	6.1	708	12	US-10-174-569-584	Sequence 584, App
26	128	6.1	708	12	US-10-174-583-584	Sequence 584, App
27	128	6.1	708	12	US-10-174-587-584	Sequence 584, App
28	128	6.1	708	12	US-10-174-589-584	Sequence 584, App
29	128	6.1	708	12	US-10-174-591-584	Sequence 584, App
30	128	6.1	708	12	US-10-175-736-584	Sequence 584, App
31	128	6.1	708	12	US-10-175-742-584	Sequence 584, App
32	128	6.1	708	12	US-10-175-744-584	Sequence 584, App
33	128	6.1	708	12	US-10-175-745-584	Sequence 584, App
34	128	6.1	708	12	US-10-175-748-584	Sequence 584, App
35	128	6.1	708	12	US-10-175-751-584	Sequence 584, App
36	128	6.1	708	12	US-10-175-754-584	Sequence 584, App
37	128	6.1	708	12	US-10-176-480-584	Sequence 584, App
38	128	6.1	708	12	US-10-176-489-584	Sequence 584, App
39	128	6.1	708	12	US-10-176-754-584	Sequence 584, App
40	128	6.1	708	12	US-10-176-755-584	Sequence 584, App
41	128	6.1	708	12	US-10-176-920-584	Sequence 584, App
42	128	6.1	708	12	US-10-176-922-584	Sequence 584, App
43	128	6.1	708	12	US-10-176-924-584	Sequence 584, App
44	128	6.1	708	12	US-10-176-984-584	Sequence 584, App
45	128	6.1	708	12	US-10-176-984-584	Sequence 584, App

#### ALIGNMENTS

RESULT 1  
US-08-755-235-4  
; Sequence 4, Application US/08755235  
; Publication No. US20030059423A1  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Wu, Jun  
; TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES  
; FILE REFERENCE: 0575/50159  
; CURRENT APPLICATION NUMBER: US/08/755,235  
; CURRENT FILING DATE: 1996-11-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Human  
US-08-755-235-4

Query Match 6.7%; Score 140; DB 8; Length 405;  
Best Local Similarity 23.7%; Pred. No. 0.001;  
Matches 78; Conservative 34; Mismatches 93; Indels 124; Gaps 16;

Qy	64	EPIITNDRFTS	-----QRYDQGGNFT--SEMIHNHVPDSGNTR-----CSLQNSRLHGS	112
Db	162	KPLVNEKGVSVKEQTRHPETGLFTLQSELM---VTPARGGDRPTFTSCSFSPGLPRHR	218	
Qy	113	AYLTQVQVGBLFIPT-----SVNLVVAENP-----CEVCLPSHWTLPDIS	154	
Db	219	ALRTAPIQPRVWEPVPLEVQLVV---EPEGGAPOGTVLTCEVPAQPS-----POLH	270	
Qy	155	WEIGLLVSHSYFYVPEPSDLQSAVSIALTPQSNGLTTCVATWKLKARKS-ATVNLTV	213	
Db	271	WMKD-----GVPLPLPPSVLLIPEIGPQDQGTQVSCVATHSHGQPSRAVVSIS	321	
Qy	214	IRCPQDTGGGINIPGVLSLSPLSGFSLPTWGVKLGLAGTWMLT-----PTCLTIRC	266	



```

XX WO9943852-A1.
EN XX
XX 02-SEP-1999.
XX PD
XX 18-FEB-1999; 99WO-KR00078.
XX PF
XX 26-FEB-1998; 98KR-0006278.
XX PR
XX (JIND/) JIN D K.
XX FA
XX (SMSU ) SAMSUNG FINE CHEM CO LTD.
XX PA
XX
XX PI Jin DK;
XX
XX WPI; 1999-527634/44.
XX DR
XX
XX Diagnosis of spinocerebellar ataxia type III (SCA III) syndrome
XX using techniques which ensure highly accurate diagnosis
XX
XX Claim 1; Page 12-13; 28pp; English.
XX
XX The invention relates to the diagnosis of spinocerebellar ataxia type
XX III (SCAIII) syndrome using reverse dot hybridisation or PCR-microplate
XX hybridisation (PCR-MPH). The method comprises attaching a portion of the
XX SCAIII gene containing 73 copies of the trinucleotide (CAG) repeat unit
XX (the present sequence) to a substrate, and hybridising with amplified
XX testee genomic DNA containing copies of the trinucleotide units, PCR
XX amplified with labeled primers (AAX89889-90). The new method is useful
XX for diagnosis of SCAIII syndrome, and for determining the severity of
XX the disease. The present sequence represents the SCAIII gene fragment
XX containing 73 trinucleotide (TNR) repeats.
XX
XX Sequence 397 BP; 124 A; 104 C; 104 G; 65 T; 0 other;
XX
XX Query Match 4.4%; Score 51.8; DB 20; Length 397;
XX Best Local Similarity 58.9%; Pred. No. 0.00039;
XX Matches 89; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
Qy 801 GCTGCTGCTGCTGCGCGCGTGGTTGTTGGCTGCACTGCTGCTGCGTGTGTTCT 860
Db 200 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 141
Qy 861 GCTCTAGAGAAAAGAGGATTCGTATTCATTTCAAAAAGAAATCTGAAAAAGAGAAGA 920
Db 140 GCTGTTGCTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 81
Qy 921 CAACAAGAACTGAGACAGAAAGTCGAAA 951
Db 80 AACTTAAAGATAAATACCATGAGAAAA 50
XX
RESULT 15
ABK10240/C
ID ABK10240 standard; DNA; 403 BP.
XX
XX AC ABK10240;
XX
XX 20-MAY-2002 (first entry)
XX
XX Trinucleotide repeat sequence #3.
XX
XX Trinucleotide repeat; fragile X syndrome; ds;
XX spinocerebellar ataxia type III; DRPLA; neuropsychiatric;
XX hereditary hypertrophic cardiomyopathy; Marfan syndrome;
XX dentatorubral and pallidoluysian atrophy;
XX spinocerebellar ataxia; X-linked spinobular atrophy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX repeat_region 152..347
XX /*tag= a
XX /rpt_type= TANDEM

```

```

FT repeat_unit      152..154
FT /*tag= b
FT /note= "CAG type repeat"
OS Homo sapiens.
XX KR2000003004-A.
XX PD 15-JAN-2000.
XX PF 25-JUN-1998; 98KR-0024064.
XX PR 25-JUN-1998; 98KR-0024064.
XX PA (SMSU ) SAMSUNG MEDICAL CENT.
XX PJ (JIND/) JIN D G.
XX PI Jin DG;
XX DR WPI; 2000-662424/64.
XX FT Plasmid vector arraying neuropsychiatric gene more than copy number
XX PT containing trinucleotide repeats its use thereof -
XX PS Disclosure; Page 12; 23pp; Korean.
XX CC This invention relates to the use of a plasmid vector array
CC containing a neuropsychiatric gene containing trinucleotide repeats.
CC Tri-nucleotide repeats and their expansion are known to be involved
CC in a number of diseases including fragile X syndrome, spinocerebellar
CC ataxia type III, hereditary hypotrophic, cardiomyopathy, dentatorubral
CC and pallidoluysian atrophy (DRPLA), spinocerebellar ataxia, Marfan
CC syndrome, X-linked spinobular atrophy. The present showing the
CC represents a nucleotide sequence of the invention showing the
CC the CAG tri-nucleotide repeats.
SQ Sequence 403 BP; 121 A; 108 C; 105 G; 69 T; 0 other;

    Query Match          4.4%; Score 51.8; DB 21; Length 403;
    Best Local Similarity 58.9%; Pred. NO. 0.0004;
    Matches 89; Conservative 0; Mismatches 62; Indels 0; Gaps 0

QY      801   GCTGTCGTGTCGGCGCGTGGTTGTTGGTGCAACTGCTGCATGCTGCGCGTGTGTTGTTCT 860
Db      215   GCTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGT 156
QY      861   GCTGTAGAAGAAAAGAGGATTTCGTAATTCATAATTTCAAAGAAATCTGAAGAAAGAGA 920
Db      155   GCTGTTGTCGTTTTTGTGTCGTGTCGTAACATTCAAAAAGTGAAGTATATTTAAAAACA 96
QY      921   CAACAAAGAAACTGAGACAGAAAGTGGAAA 951
Db      95   AACTTAAAGNATAAATACACCATGAGNAAA 65

Search completed: November 12, 2003, 17:05:22
Job time : 347.543 secs

```

Search completed: November 12, 2003, 17:05:22  
Job time : 347.543 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:41:25 ; Search time 2799.55 Seconds  
(without alignments)  
10140.072 Million cell updates/sec

Title: US-09-729-264-3

Perfect score: 1168

Sequence: 1 agtgatcatgtggcaggag.....gtaatacaactgtagtatatg 1168

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estio.\*

8: em\_hic.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	781.4	66.9	1201	13 BX399881	BX399881 BX399881
2	538.8	46.1	941	13 BX403420	BX403420 BX403420
3	509.2	43.6	917	13 BX370118	BX370118 BX370118
4	498.2	42.7	1201	13 BX358190	BX358190 BX358190

5	456	39.0	970	13 BX370209	BX370209 BX370209
6	326	27.9	1510	11 AK086973	AK086973 Mus muscu
7	306	26.2	784	12 BG863468	BG863468 602796941
8	276.4	23.7	794	10 BF784177	BF784177 602108039
9	267.4	22.9	725	10 BG740428	BG740428 602633817
10	266.4	22.8	1148	13 BX370210	BX370210 BX370210
11	256.8	22.0	631	10 BB663870	BB663870 BB663870
12	242	20.7	527	10 BE032610	BE032610 BX370210
13	226.2	19.4	1201	13 BX358189	BX358189 BX358189
14	204.8	17.5	556	4 BX511978	BX511978 RZPD Mus
15	193.2	16.5	517	10 BA476432	BA476432 159670 BA
16	190.8	16.3	622	29 AG069679	AG069679 Pan trogl
17	190.8	16.3	677	29 AG107877	AG107877 Pan trogl
18	185.8	15.9	524	9 AV601192	AV601192 AV601192
19	181.8	15.6	471	9 AA265274	AA265274 mx91d09.r
20	181.8	15.6	611	10 AW990468	AW990468 uf37d05.y
21	181.8	15.6	691	14 BY708046	BY708046 BY708046
22	181.8	15.6	931	10 BG173684	BG173684 602336823
23	181.8	15.6	1552	11 AK008060	AK008060 Mus muscu
24	181.8	15.6	1552	11 AK078934	AK078934 Mus muscu
25	180.8	15.5	486	9 AI425363	AI425363 mx91d09.y
26	180.2	15.4	663	10 BE376590	BE376590 601226649
27	178.6	15.3	650	10 BB625217	BB625217 BB625217
28	165	14.1	600	28 AZ379623	AZ379623 IM0134622
29	153.2	13.1	216	10 BG206666	BG206666 RST36117
30	152.4	13.0	388	14 CB805443	CB805443 AMGNNUC:C
31	149	12.8	556	9 AI788300	AI788300 uk56d01.y
32	147.4	12.6	530	9 AI790785	AI790785 uk28a12.y
33	145.4	12.4	291	28 AZ411779	AZ411779 IM0185804
34	141.6	12.1	423	10 BB846577	BB846577 BB846577
35	129.8	11.1	426	13 BY033616	BY033616 BY033616
36	128.8	11.0	471	10 BF040046	BF040046 BP250004B
37	126.4	10.8	351	13 BY100408	BY100408 BY100408
38	124	10.6	685	29 AG142221	AG142221 Pan trogl
39	123.2	10.5	394	10 BB846133	BB846133 BB846133
40	118.4	10.1	346	13 BY143304	BY143304 BY143304
41	117.4	10.1	369	13 BY036814	BY036814 BY036814
42	110	9.4	785	12 BI452873	BI452873 603170211
43	102.8	8.8	889	12 BI454276	BI454276 603171509
44	100	8.6	185	14 N47851	N47851 YW95h05.r1
45	100	8.6	234	14 N93995	N93995 za66f09.r1

#### ALIGNMENTS

RESULT 1  
BX399881  
LOCUS BX399881 1201 bp mRNA linear EST 13-MAY-2003  
DEFINITION BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0DI085YA24 5-PRIME, mRNA sequence.  
ACCESSION BX399881  
VERSION BX399881.1 GI:30622019  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequenage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10299.f, and  
it belongs to a clone representative of this cluster. For more  
information about this cluster and the virtual cDNA, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DI085BAL2QP1&cluster=10299.f. Contact :  
Feng liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0D1085B112Q1.  
Location/Qualifiers  
source  
1. 1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1085YA24"  
/tissue type="PLACENTA COT 25-NORMALIZED"  
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 290 a 285 c 301 g 297 t 28 others  
ORIGIN  
Query Match 66.9%; Score 781.4; DB 13; Length 1201;  
Best Local Similarity 93.1%; Pred. No. 1.2e-131;  
Matches 835; Conservative 21; Mismatches 37; Indels 4; Gaps 4;  
QY 45 GTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCAAAATGCAAGAGTCTCTGAAGGCT 104  
DB 272 GTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCAAAATGCAAGAGTCTCTGAAGGCT 331  
QY 105 CCCAGGCTCGTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 164  
DB 332 CCCAGGCTCGTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 391  
QY 165 GTGACATGGTGTGTAAGCGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCTTCA 224  
DB 392 GTGACATGGTGTGTAAGCGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCTTCA 451  
QY 225 CCTCTCAGAGGTACACACAGCGCGGAACTTCCACTCGGAGATGATCATCCACAAATGTGG 284  
DB 452 CCTCTCAGAGGTACACACAGCGCGGAACTTCCACTCGGAGATGATCATCCACAAATGTGG 511  
QY 285 AGCCAGTCAATTCGGGGAACATCAGATGAGCGCTCCAGACAGTCCCTGCATGATCTG 344  
DB 512 AGCCAGTCAATTCGGGGAACATCAGATGAGCGCTCCAGACAGTCCCTGCATGATCTG 571  
QY 345 CTTACTCTTACCGTCCAAAGTATGGAGAGCTGTTTCATTCACAGTGTAACTTTGTAGTCG 404  
DB 572 CTTACTCTTACCGTCCAAAGTATGGAGAGCTGTTTCATTCACAGTGTAACTTTGTAGTCG 631  
QY 405 CTGAGAAATGAACCTTGTGAAGTACTTGTCTACCTCACACTGGACCTGGTCCCGGATA 464  
DB 632 CTGAGAAATGAACCTTGTGAAGTACTTGTCTACCTCACACTGGACCTGGTCCCGGATA 691  
QY 465 TTTCTGGGAGCTGGTCTCCTGGTCAGCCATTCAGCTATATTTTGTTCGGAGCCCA 524  
DB 692 TTTCTGGGAGCTGGTCTCCTGGTCAGCCATTCAGCTATATTTTGTTCGGAGCCCA 751  
QY 525 GCGACCTTCAAAGTGCAGTGCATCTCGGTCTGCACCCCAAGCAATGGACTTTGA 584  
DB 752 GCGACCTTCAAAGTGCAGTGCATCTCGGTCTGCACCCCAAGCAATGGACTTTGA 811  
QY 585 CTTGGTGGCTACCTGGAGAGCTTGAAGCCCGCAAGTCTGCAACTGTAATCTCACTG 644  
DB 812 CTTGGTGGCTACCTGGAGAGCTTGAAGCCCGCAAGTCTGCAACTGTAATCTCACTG 871  
QY 645 TGATTTCGGTGTCCCAAGACACTGGAGGTGGTATTAATATTCAGAGTGTATTCAGATT 704  
DB 872 TGATTTCGGTGT-CCCAAGACACTGGAGGTGGTATTAATATTCAGAGTGTATTCAGATT 929  
QY 705 TACCGAGTTTAGTTTTCATTGCCCTACTTGGGGCAAGTTGGACTTGGACTAGCAGCA 764  
DB 930 TACCGAGTTTAGTTTTCATTGCCCTACTTGGGGCAAGTTGGACTTGGACTAGCAGCA 989  
QY 765 CCAATGCTTCTGAGCCGAGGTGATCTTTCATTAATACGCTGTGCTGCTGCCGCTCGTT 824  
DB 990 CCAATGCTTCTGAGCCGAGGTGATCTTTCATTAATACGCTGTGCTGCTGCCGCTCGTT 1049

QY 825 GTTGTGGCTGCAACTGCTGCTGCGTGTGTTTCTGCTAGAGAAAAAGAGATTTC 884  
DB 1050 GATGKTGCTGCAACT-CTGCKRCGGTAKTTGTTCTTCTGTAGAA-RAAARARGGTTYS 1107  
QY 885 GTATTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAAACTGAGACAG 941  
DB 1108 GTATCMATTTTCAAAARAWATCTGAAAAGAGAAAGACAAMAAARACACGAAACAGMAGWG 1164  
RESULT 2  
LOCUS BX403420 941 bp mRNA linear EST 13-MAY-2003  
DEFINITION BX403420 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1085YA24 5-PRIME, mRNA sequence.  
ACCESSION BX403420  
VERSION BX403420.1 GI:30607302  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 941)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
10299.f, and it belongs to a clone representative of this cluster.  
For more information about this cluster and the virtual cDNA, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS5AA017ZE02RM1&cluster=10299.f. Contact :  
Peng Liang Email : fliang@lifetech.com URL : Genitrogen Corporation 1600  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS5AA017ZE02RM1.  
Location/Qualifiers  
source  
1. 941  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1085YA24"  
/tissue type="PLACENTA COT 25-NORMALIZED"  
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 214 a 241 c 255 g 227 t 4 others  
ORIGIN  
Query Match 46.1%; Score 538.8; DB 13; Length 941;  
Best Local Similarity 98.4%; Pred. No. 1e-87;  
Matches 565; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
QY 45 GTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCAAAATGCAAGAGTCTCTGAAGGCT 104  
DB 286 GTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCAAAATGCAAGAGTCTCTGAAGGCT 345  
QY 105 CCCAGGCTCGTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 164  
DB 346 CCCAGGCTCGTTCAACTGCAACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 405  
QY 165 GTGACATGGTGTGTAAGCGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCTTCA 224  
DB 406 GTGACATGGTGTGTAAGCGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCTTCA 465  
QY 225 CCTCTCAGAGGTACGACCCAGCGCGGAACTTCCACTCGGAGATGATCATCCACAAATGTGG 284  
DB 466 CCTCTCAGAGGTACGACCCAGCGCGGAACTTCCACTCGGAGATGATCATCCACAAATGTGG 525

QY	285	AGCCAGTGATTCCGGGAGCAATCAGATGCAGCCTCCAGAACAGTGCCTGCATGATCTG	344					
Db	526	AGCCAGTGATTCCGGGAGCAATCAGATGCAGCCTCCAGAACAGTGCCTGCATGATCTG	585					
QY	345	CTTACCTTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCCCAAGTCTTAATCTTTAGTCG	404					
Db	586	CTTACCTTTACCGTCCAAAGTTATGGAGAGCTGTTTCATTCCCAAGTCTTAATCTTTAGTCG	645					
QY	405	CTGAGAATGAACCTTGTGAAGTTACTTGTACCTCTACATCTGGAACCTGGCTCCCGGATA	464					
Db	646	CTGAGAATGAACCTTGTGAAGTTACTTGTACCTCTACATCTGGAACCTGGCTCCCGGATA	705					
QY	465	TTTCCTGGAGCTCGGTCTCCTGTCAGCCATTCAAGCTATTATTTTG-TTCCGGAGCCC	523					
Db	706	TTTCCTGGAGCTCGGTCTCCTGTCAGCCATTCAAGCTATTATTTTG-TTTCGGAGCCC	765					
QY	524	AGCGACCTTCAAAGTGCAGTGCAGCATCTCGCTCTGACCCACAGACGAATGGGACTTTG	583					
Db	766	AGCGACCTTCAAAGTGCAGTGCAGCATCTCGCTCTGACCC-CAGACGAATGGGACTTTG	824					
QY	584	ACTTGGCTGGTACTCTGGAAGAGCCTGAAGCCC	617					
Db	825	ACTTGGCTGGTCTCCTGGAAGAGCCTGAAGCCC	858					
RESULT	3	917 bp mRNA linear EST 08-MAY-2008						
LOCUS	BX370118							
DEFINITION	BX370118 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA							
ACCESSION	Clone CSODI076YN22 5-PRIME, mRNA sequence.							
VERSION	BX370118							
KEYWORDS	EST.							
SOURCE	Human							
ORGANISM	Homo sapiens (human)							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
TITLE	1 (bases 1 to 917)							
JOURNAL	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.							
COMMENT	Full-length cDNA libraries and normalization							
Unpublished								
Contact: Genoscope								
Genoscope - Centre National de Sequencage								
BP 191 91006 EVRY cedex - France								
Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr								
Library was constructed by Life Technologies, a division of								
Invitrogen. This sequence belongs to sequence cluster 2250.r For								
more information about this cluster, see								
http://www.genoscope.cns.fr/								
cgi-bin/cluster.cgi?sed=CSOBAG037ZE05_CS03493_2&cluster=2250.r.								
Contact : Feng Liang Email : fliang@lifetech.com URL :								
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600								
Faraday Avenue Genoscope sequence ID : CSOBAG037ZE05_CS03493_2.								
Location/Qualifiers								
FEATURES	1..917							
source	/organism="Homo sapiens"							
	/mol_type="mRNA"							
	/db_xref="taxon:9606"							
	/clone="CSODI076YN22"							
	/tissue_type="PLACENTA COT 25-NORMALIZED"							
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"							
	/note="1st strand cDNA was primed with a NotI-oligo(dT)							
	primer. Five prime end enriched, double-strand cDNA was							
	digested with Not I and cloned into the Not I and BcoR v							
	sites of the pCMVSPORT 6 vector. Library was normalized."							
BASE COUNT	272 a 203 c 190 g 252 t							
ORIGIN								
Query Match 43.6%; Score 509.2; DB 13; Length 917;								
Best Local Similarity 99.1%; Pred. No. 2.3e-82;								
Matches 523; Conservative 0; Mismatches 3; Indels 2; Gaps 1								
QY	641	ACTGTGATTCGGTCTCCCAAGACACTGGAGGTGGTATTATATATCCAGGTGATTATCA	700					



99279253 MEDLINE PUBMED REFERENCE AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374 MEDLINE PUBMED REFERENCE AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
20530913 MEDLINE PUBMED REFERENCE AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660 MEDLINE PUBMED REFERENCE AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 1510)	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222).
99279253 MEDLINE PUBMED REFERENCE AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374 MEDLINE PUBMED REFERENCE AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
20530913 MEDLINE PUBMED REFERENCE AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660 MEDLINE PUBMED REFERENCE AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 1510)	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222).

99279253 MEDLINE PUBMED REFERENCE AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374 MEDLINE PUBMED REFERENCE AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
20530913 MEDLINE PUBMED REFERENCE AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660 MEDLINE PUBMED REFERENCE AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 1510)	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222).

QY 701 AGTTTACGAGTTTAGGTTTTTCATTGCTACTTTGGGCAAAAGTTGGACTTGACGA 760  
 Db 799 -----AGGCCAGCACTGCCGACCTGGGCCATCATCTCTGCGCAGTGCC 844  
 QY 761 GGCACCATGCTCTGACGCCGACGTGTAATCTTACAATACGCTGCTGCTGCCCGCT 820  
 Db 845 TTTTCCCTTGCCTTGATCCTGATCAATGTTTGTGATTAATTAATCTGTGCTGTGT--- 901  
 QY 821 CGTTGTTGGCTGCAACTGCTGCTGCCGTTGTTTCTGCTGTAGAAGAAAAGAGGA 880  
 Db 902 -----GCTCCAGGAGAAAGGAAGAACTACTTATCAAAATGAAATAAGAA 951  
 QY 881 TTTCTATTCAATTTCAAAAGAAATCTGAAAGAGAGACAAACAAAGAACTGAGACA 940  
 Db 952 ATCTGCAAAACATGAGGACAA-----ACAAAGCAGATCCGGAGACAAAGTTA 997  
 QY 941 GAAGTGAAGATGAATCTCGGCTACAAATCCAGATGATCCTGAACAAAGAAACAGTAGC 1000  
 Db 998 AAAAGTGAAGAGGAAACTACGGGTACAGTTTCGGATGAGGCAAAAGGCTGCACAGACTGA 1057  
 QY 1001 TCTCTCCCTCCCAATCTGTGAATCCAGTATCCTGAACAAAGAAACAGTAGC 1054  
 Db 1058 TCTCTCCCTCCTAAATCTGCTGAAGTCAAGCTTCCAGAAAAACGACAGTAGC 1111

RESULT 7  
 BG863468 784 bp mRNA linear EST 29-MAY-2001  
 LOCUS 602796941F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4917803 5',  
 DEFINITION mRNA sequence.

ACCESSION BG863468  
 VERSION BG863468.1 GI:14214006  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 JOURNAL 1 (bases 1 to 784)  
 COMMENT NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth  
 Ph.D.

cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

Plate: LLAM10830 row: b column: 12  
 High quality sequence stop: 768.

FEATURES  
 Location/Qualifiers

1..784  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NMRI"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4917803"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP Mam4"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Priscilla Furth,  
 NIH Reference for transgenic model: Li et al., Cell Growth  
 and Differentiation 7, 3-11 (1996)."  
 BASE COUNT 188 a 202 c 208 g 186 t  
 ORIGIN

Query Match 26.2%; Score 306; DB 12; Length 784;  
 Best Local Similarity 70.6%; Pred. No. 1.4e-45;  
 Matches 436; Conservative 0; Mismatches 180; Indels 2; Gaps 2;

QY 42 CCGGTTCTGGGCTGTAATGAAGTCATAGAGGCCCCCAAAATGCAGAGTTCCTGGAAGG 101  
 Db 148 CAGCTTCGGATCCAGTTATCAGATCATAGAAAGTTCCTCAGAATGTACAGTCTCTAAAGG 207  
 QY 102 GCTCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTC 161  
 Db 208 ACTCAGAGGCTCACTTCAACTGCACCGTGAATCAGCGCTGGAAGCTTCTCATGTGGACTC 267  
 QY 162 TCAGTGACATGGTGGTCTAAGCGTCAAGCCCATAGGACCCCATCATCAACATGACCCCT 221  
 Db 268 TTAACCAAAATGGTGGTGTGAGTCTCACCAACCAAGGACCCCATCATCAACCAACACCGCT 327  
 QY 222 TCACCTCTCAGAGGTAGACACAGGGCGGAATCTTCACTCGAGATGATCATCCACAATG 281  
 Db 328 TCACCTATGCCAGTTTACAAACAGCACTGACAGCTTCATCTCGAGTTTGATCATCCATGATG 387  
 QY 282 TGGAGCCCAAGTATTCGGGGAACATCAGATGCAGCTCCAGAACAGACAGTCGCCCTGCATGGAT 341  
 Db 388 TGCAGCCCAAGTACTCGGGATCCGTCGAATGCAGCTGCAGAACAGCCATGCGTTTGGAT 447  
 QY 342 CTGCTTACCTTACCGTCCAAAGTTATGGAGAGAGCTGTTCAATTCCTCCAGTGTTAATCTGTAG 401  
 Db 448 CTGCTTCTCTCATGTGCAAGTCATGCGGACCCCTGAACATTCCTAGCAACACCTTTATAG 507  
 QY 402 TCGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCTGGGCTCCCGG 461  
 Db 508 TCACAGGGGTGAACCTGTGAATGTGACTTGTATGCGGTGGGTGGAGCTCACTCCCGG 567  
 QY 462 ATATTTCTGGAGCTCGGCTCTCTCTGTCAGCCATTCAGCTATTAATTTGTTCCGAGC 521  
 Db 568 ATATTTCTGGAGCTTGAGGTTCGCGTAAGCCATTCAGTTACAATTCCTTCTGGAGC 627  
 QY 522 CCAGGACCTTCAAGTGCAGTGAAGCATCTCGGCTCTGACCCACAGAGCAATGGGACTT 581  
 Db 628 CGGCAACTTTTATGAGGCTCTTGAGTGTCTTGAGCTGCAGCTCACAACA-TGGGCAACGGGACCT 686  
 QY 582 TGACTTGGTGGCTACCTGGAAGAGCCTGAAGGCCGCAAGTCTGCAACTCTCAATCTCA 641  
 Db 687 TGACTTGTGGCAGAGCTGAAGACTTGCAGGCAGCAAGTCTTAACTGTCAA-CTGA 745  
 QY 642 CTGTGATTCGGTGTCCCC 659  
 Db 746 CTGTGTTTCAGCTTCCAC 763

RESULT 8

BF784177

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

BF784177

VERSION

BF784177.1

KEYWORDS

EST.

SOURCE

Mus musculus

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 794)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9844 row: n column: 10

High quality sequence stop: 694.

FEATURES

source  
1. 794  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4236489"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Kid14"  
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library." |

BASE COUNT 186 a 214 c 215 g 179 t

ORIGIN

Query Match 23.7%; Score 276.4; DB 10; Length 794;  
Best Local Similarity 69.3%; Pred. No. 3.2e-40;  
Matches 405; Conservative 0; Mismatches 176; Indels 3; Gaps 2;

QY 42 CCGGTTCTGGGTCTGTAATCAAGTCTATAGAGGCCCCCAAAATGCGAGAGTCTGAGG 101  
Db 148 CAGCTCCGGATCCAGTTATCAGATCATAGAGGTCTCTAGAATGTAACTGCTTAAAGG 207  
QY 102 GCTCCAGGCTCGTCTCAACTGCACCGCTCTCCAGGCTGGAAGCTCATCATGTTGGCTC 161  
Db 208 ACTCAGAGGCTCACTTCAACTGCACCGTGACTCAGGCTGGAAGCTTCTCATGTGACTC 267  
QY 162 TCAGTGACATGGTGTCTGAAGCGTCAAGCGCCATGAGAGCCCATCATCACCAGTACCGCT 221  
Db 268 TTAACCAATAGTGTCTGAGTCTCACCACCCAGGACCCATCATCACCAGTACCGCT 327  
QY 222 TCACCTCTCAGAGTACGACGAGGCGGAATTCACCTCGGAGATCATCATCCACATG 281  
Db 328 TCACCTTAGCGATTACAGAGCACTGACAGCTTCATCTCGAGTTGATCATCTGATG 387  
QY 282 TGAGGCCAGTGATTCGGGGAACATCAGATGACGCTCCAGAAAGTCCGCTGATGGAT 341  
Db 388 TGAGGCCAGTGATTCGGGATCCGTGCAATGACGCTCGAGAACGCTATGGTGGAT 447  
QY 342 CTGCTTACCTTACCGTCCAGTTATGGGAGAGCTGTTTCATTCCAGTGTAACTTGTAG 401  
Db 448 CTGCTTCTCTCAGTGCAGATCATGGGGACCTTGAAACATTCCTAGCAACACCTTATAG 507  
QY 402 TCCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGACCTGGCTCCCGG 461  
Db 508 TCAGTGGGTGAACCTTGAATGTGACTTGTATGCTGATGCGGCTGGACCTCCTCCCGG 567  
QY 462 ATATTCTCTGGAGCTCGGTCTCTCTGTCAGCCATCAAGCTATATTATTTGTCGGAGC 521  
Db 568 ATATTCTCTGGAGCTTGAGGTTCCCGTAAAGCCATTCGAGTTACAAATCTCTTCTGGAGC 627  
QY 522 CCAGGACCTTCAAGTGCAGTGCATCTGGCTCTGACCCACAGAGCAATGGGACTT 581  
Db 628 C--GGGCACCTTATGAGGGTCTGAGTGTCTGGAACCTCA-CACACTCGGCAACGGGACCT 684  
QY 582 TGACTTTCGTTGGTACCTGGAAGAGCTGGAAGGCCCGCAAGTCT 625  
Db 685 TGACTTGTGGCAGAGCTGAGGACTTGCAGCCAGCAGTCT 728

RESULT 9

BG740428

LOCUS

DEFINITION

602633817F1 NCI\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4778789 5',

mRNA sequence.

ACCESSION

BG740428

VERSION

BG740428.1 GI:14051081

KEYWORDS

BG740428 725 bp mRNA linear EST 15-MAY-2001  
EST.  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 725)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: gspbs-r@mail.nih.gov  
Tissue Procurement: James Cleaver, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA  
Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10635 row: b column: 06  
High quality sequence stop: 725.

FEATURES

source

1. 725

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4778789"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_Skn3"

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.5kb. Library constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library." |

BASE COUNT 227 a 163 c 159 g 176 t

ORIGIN

Query Match 22.9%; Score 267.4; DB 10; Length 725;

Best Local Similarity 99.6%; Pred. No. 1.4e-38;

Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 899 AAGAAATCTGAAAAGAGAGACAAACAAAGAACTCAGACAGAAATGGAATGAAAC 958

Db 377 AGGAAATCTGAAAAGAGAGACAAACAAAGAACTCAGACAGAAATGGAATGAAAC 436

QY 959 TCCGGCTACATTCAGATGAACAAAGAACACACAGACCGCTTCTCTCCCTCCCAATCC 1018

Db 437 TCCGGCTACATTCAGATGAACAAAGAACACACAGACCGCTTCTCTCCCTCCCAATCC 496

QY 1019 TGTGAATCCAGTGATCTGTAACAAAGAAACAGTAGCTGTGCGCTCTCTCCCTCCCAATCC 1078

Db 497 TGTGAATCCAGTGATCTGTAACAAAGAAACAGTAGCTGTGCGCTCTCTCCCTCCCAATCC 556

QY 1079 GATCAACGTCACCCAGCCAGCAAGTCAATCCAGAGCTTTTAAATCTGGCCAGTCT 1138

Db 557 GATCAACGTCACCCAGCCAGCAAGTCAATCCAGAGCTTTTAAATCTGGCCAGTCT 616

QY 1139 GAGAGGTGAGTAACTCAACTGTAGTATA 1167

Db 617 GAGAGGTGAGTAACTCAACTGTAGTATA 645

RESULT 10

BX370210

LOCUS

DEFINITION

BX370210 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CSODI085YA24 5-PRIME, mRNA sequence.

ACCESSION

BX370210

VERSION

BX370210.1 GI:30451881

KEYWORDS

EST.

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1148)

AUTHORS



```
Matches 336; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 42 CCGGTTCTGGGTCTGGTAATGAAGTCAATAGAGCCGCCAAAATGGAAGTCTGTAAGG 101
Db 161 CAGCTTCGGGATCCAGTTATCAGATTCATAGAGGTCCTCAGAATGTAAACATGCTTAAAGG 220
QY 102 GCTCCCAAGGTCGGTTCACACTGCACCGCTCTCCAGGCTGGAAGCTCATCATGTGGGCTC 161
Db 221 ACTCAGAGGCTCACTTCACTGACCGTGAAGTCACTCAGGCTGGAAGCTTCTCATGTGGACTC 280
QY 162 TCAGTGCATCGTGGTCTTAAGCGTCAAGCGCTCAGGCGCCATCGAGCCCATCATCACCNAATGACCGCT 221
Db 281 TTAACCAAAATGGTGGTCTGAGTCTCAGTCTCACCCAGGACCCATCATCACCNAACACCGCT 340
QY 222 TCACCTCTCAGAGTACAGACAGGCGGGAAGTCTCACTCGGAGATGATCATCACAATG 281
Db 341 TCACCTATGCGCAGTTACACAGCACTGACAGCTTCACTCGGAGTTGATCATCATCATGATG 400
QY 282 TGGAGCCCAAGTATTCGGGGAAACATCAGATGACGCTCCAGAACAGTGCCTGTCATGGAT 341
Db 401 TGCAGCCCAAGTACTCGGATCCGTCGAATGCGAGCTTCAGAACAGCGATGGTGGAT 460
QY 342 CTGCTTACCTTACCGTCCAGTTATGGGAGAGCTGTTTCATTCACAGTGTAAATCTTGTAG 401
Db 461 CTGCTTTCCTCTCAGTGCAGATCATGGGGACCCCTGAACATTCCTAGCAACAACCTTATAG 520
QY 402 TCGTGAAGTGAACCTTGTGAAGTACTTGTCTACCTCAGCTGACCTGCGCTCCCGG 461
Db 521 TCACTGAGGTGAACCTGTGAATGTGACTTGTATGCGGTGGGCTGAGCTCACTTCCCGG 580
QY 462 ATATTTCTCGGAGCTCGGTCTCTCTGTCAGCCATTCAGCTATTAATT 509
Db 581 ATATTTCTCGGAGCTTGAAGTTCCCGTGAAGCAATTCGAGTTACAATT 628

RESULT 12
BE032610
LOCUS
DEFINITION 132035 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE032610
VERSION BE032610.1 GI:8327619
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 527)
Fahrenkrug, S.C., Smith, T.P.L., Preking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R., Quackenbush,
J., and Keefe, J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 63 row: E column: 13
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..527
/organism="Sus scrofa"
/mol_type="mRNA"
```

```
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1PIG"
/notes=Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 105 a 147 c 176 g 99 t
ORIGIN
Query Match 20.7%; Score 242; DB 10; Length 527;
Best Local Similarity 78.4%; Pred. No. 5.6e-34;
Matches 290; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 42 CCGGTTCTGGGTCTGGTAATGAAGTCAATAGAGCCGCCAAAATGGAAGTCTGTAAGG 101
Db 158 CAGCTTCGGATCCGGCAATGMAATCTTAGAGGTCCTCAGGAACGTCACGGCGGTGAAGG 217
QY 102 GCTCCCAAGGTCGGTTCACACTGCACCGCTCTCCAGGCTGGAAGCTCATCATGTGGGCTC 161
Db 218 GCTCGGAGGCTCGCTTCACTGACCACTCTCCAGGCTGGAAGCTCATCATGTGGGCTC 277
QY 162 TCAGTGCATCGTGGTCTTAAGCGTCAAGCGCTCAGGCGCCATCGAGCCCATCATCACCNAATGACCGCT 221
Db 278 TGAAGGCAACGGTGGTCTGAGCGTCAAGCGCCACCGAGCCCATCATCACCNAACCGCAT 337
QY 222 TCACCTCTCAGAGTACAGACAGGCGGGAAGTCTTCACTCGGAGATGATCATCACAATG 281
Db 338 TCAGCTTGAAGTTATGAAGAGGCGGTAACTTCATCTCTGAGATGATCATCATCATGACG 397
QY 282 TGGAGCCCAAGTATCGGGGAACATCAGATGACGCTCCAGAACAGTGCCTGCGCTGATGAT 341
Db 398 TGCAGCTCAGTATGCGCGGCGAGTCAATGAGCGCTCCAGAACAGTGCAGGATGGAT 457
QY 342 CTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTTCATTCGCCAGTGTAAATCTTGTAG 401
Db 458 CCGCTTTCCTTCGTCCTCAAGTCATGGGGTGTGCTCATTCGCGGAGGAGCGCTCGTAG 517
QY 402 TCCTGTAGAA 411
Db 518 TCCTTGGCGA 527

RESULT 13
BX358189/c
LOCUS
DEFINITION BX358189 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI035YC01 3-PRIME, mRNA sequence.
ACCESSION BX358189
VERSION BX358189.1 GI:30370198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f
Contact : Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODI035AB01NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```



```
Db          500 CCGCCTTCCTTCTGTC 517
Search completed: November 12, 2003, 23:19:37
Job time : 2801.55 secs
```





[illegible]

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-491-356C-2

Query Match
Best Local Similarity 3.8%; Score 44.8; DB 4; Length 6794;
Matches 55; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCCGCGTGTGTGTGGCTGCAACTGCTGCTGCCGTTGTTTCT 860
DB 6349 GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6290

QY 861 GCTGTAGAGAA 872
DB 6289 GCTGCTGCCGCA 6278

RESULT 10
US-09-491-356C-1/c
; Sequence 1, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Gims, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491,356C
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 55298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (485)..(485)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (838)..(838)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (16728)..(16728)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22750)..(22750)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22756)..(22756)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (28519)..(28519)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (44804)..(44804)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (45002)..(45002)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54049)..(54049)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54226)..(54226)
; OTHER INFORMATION: n is not determined
US-09-491-356C-1

Query Match
3.8%; Score 44.8; DB 4; Length 55298;
Best Local Similarity 76.4%; Pred. No. 0.011;
Matches 55; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCCGCGTGTGTGTGGCTGCAACTGCTGCTGCCGTTGTTTCT 860
DB 6349 GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6290

QY 861 GCTGTAGAGAA 872
DB 6289 GCTGCTGCCGCA 6278

RESULT 11
US-09-620-312D-475/c
; Sequence 475, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 475
; LENGTH: 3302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(2563)
US-09-620-312D-475

Query Match
3.7%; Score 43.4; DB 4; Length 3302;
Best Local Similarity 69.4%; Pred. No. 0.018;
Matches 59; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCCGCGTGTGTGTGGCTGCAACTGCTGCTGCCGTTGTTTCT 860
DB 527 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468

QY 861 GCTGTAGAGAAAGAGAGATTTCG 885
DB 467 GCTGCTGCTGCTGCTGAGGATGACG 443

RESULT 12
US-09-086-663A-70/c
; Sequence 70, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA
```

APPLICANT: KARSENTY, GERARD  
TITLE OF INVENTION: OSF2/CBFAI COMPOSITIONS AND METHODS OF USE  
FILE REFERENCE: UTSC:525  
CURRENT APPLICATION NUMBER: US/09/086,663A  
PRIOR FILING DATE: 1998-05-29  
PRIOR APPLICATION NUMBER: 60/080,189  
PRIOR FILING DATE: 1998-03-24  
PRIOR APPLICATION NUMBER: 60/048,430  
PRIOR FILING DATE: 1997-05-29  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 70  
LENGTH: 2294  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1644)  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Primer  
US-09-086-663A-70

Query Match 3.7%; Score 43.2; DB 4; Length 2294;  
Best Local Similarity 79.7%; Pred. No. 0.017;  
Matches 51; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCCCGCGTGTGTGGCTGCAACTGCTGCTGCCGTTGTGTTCT 860  
Db |||||  
226 GCTGCTGCTGCTGCTGTTGTGTGCTGTGTTGCTGTGTTGCTGCTGCTGCTGCT 167  
QY 861 GCTG 864  
Db |||||  
166 GCTG 163

RESULT 13  
US-09-086-663A-1/c  
Sequence 1, Application US/09086663A  
Patent No. 6518063  
GENERAL INFORMATION:  
APPLICANT: DUCY, PATRICIA  
TITLE OF INVENTION: OSF2/CBFAI COMPOSITIONS AND METHODS OF USE  
FILE REFERENCE: UTSC:525  
CURRENT APPLICATION NUMBER: US/09/086,663A  
PRIOR FILING DATE: 1998-05-29  
PRIOR APPLICATION NUMBER: 60/080,189  
PRIOR FILING DATE: 1998-03-24  
PRIOR APPLICATION NUMBER: 60/048,430  
PRIOR FILING DATE: 1997-05-29  
NUMBER OF SEQ ID NOS: 83  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3334  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Primer  
US-09-086-663A-1

Query Match 3.7%; Score 43.2; DB 4; Length 3334;  
Best Local Similarity 79.7%; Pred. No. 0.021;  
Matches 51; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCCCGCGTGTGTGGCTGCAACTGCTGCTGCCGTTGTGTTCT 860  
Db |||||  
534 GCTGCTGCTGCTGCTGTTGTGCTGTGTTGCTGTGTTGCTGCTGCTGCTGCT 475  
QY 861 GCTG 864  
Db |||||  
474 GCTG 471

RESULT 14  
US-09-491-356C-7/c  
Sequence 7, Application US/09491356C  
Patent No. 6566061  
GENERAL INFORMATION:  
APPLICANT: Philibert, Robert A.  
APPLICANT: Ginns, Edward I.  
APPLICANT: Delisi, Lynn  
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13  
FILE REFERENCE: 9465.6US11  
CURRENT APPLICATION NUMBER: US/09/491,356C  
CURRENT FILING DATE: 2000-01-26  
PRIOR APPLICATION NUMBER: PCT/US99/09365  
PRIOR FILING DATE: 1999-04-29  
PRIOR APPLICATION NUMBER: 60/083,465  
PRIOR FILING DATE: 1998-04-29  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7  
LENGTH: 6558  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-491-356C-7

Query Match 3.7%; Score 43; DB 4; Length 6558;  
Best Local Similarity 77.6%; Pred. No. 0.035;  
Matches 52; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCCCGCGTGTGTGGCTGCAACTGCTGCTGCCGTTGTGTTCT 860  
Db |||||  
6181 GCTGCTGCTGCTGTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6122  
QY 861 GCTGTAG 867  
Db |||||  
6121 GCTGTTG 6115

RESULT 15  
US-08-531-927B-1/c  
Sequence 1, Application US/08531927B  
Patent No. 5840491  
GENERAL INFORMATION:  
APPLICANT: Kakizuka, Akira  
TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph  
TITLE OF INVENTION: Disease Gene and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173-4799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,927B  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP H6-251600  
FILING DATE: 21-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: ATH95-01A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240

```
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1776 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1115
US-08-531-927B-1

Query Match      3.7%; Score 42.8; DB 2; Length 1776;
Best Local Similarity 61.8%; Pred. No. 0.019;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      801 GCTGCTGCTGCGCGCTGCTGTTGTTGGGCTGCAACTGCTGCTGCGGTTGTTCTTCT 860
Db      975 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916

Qy      861 GCTGTAGAAGAAAGAGGATTTCGTATTCAATTTCAAAAGAAATCTGAA 910
Db      915 GCTGCTGTTTTCAAGTAGGCTTCTCGTCTCTTCGAGACTTCTTGAA 866
```

Search completed: November 12, 2003, 23:23:37  
Job time : 78.2582 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 20:56:43 ; Search time 391.507 Seconds  
(without alignments)  
9519.735 Million cell updates/sec

Title: US-09-729-264-3

Perfect score: 1168

Sequence: 1 agtatacgtgagcaggag.....gtaatacaactgtgtatag 1168

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269	23.0	474	11	US-09-918-995-3342
2	196.8	16.8	401	9	US-09-864-761-16305
3	136.2	11.7	398	10	US-09-983-965-4945
C 4	46	3.9	3577	13	US-10-008-739A-1
C 5	44.8	3.8	1310	10	US-09-849-243-13
C 6	44.8	3.8	2614	11	US-03-822-846-491
C 7	44.8	3.8	3263	10	US-03-849-243-15
C 8	44.8	3.8	4286	10	US-09-849-243-14
C 9	44.8	3.8	5085	14	US-10-198-846-9854
C 10	44.8	3.8	6004	10	US-09-880-107-1748
C 11	44.2	3.8	2790	12	US-10-029-386-22626
C 12	44	3.8	299	12	US-09-814-353-4844
C 13	44	3.8	299	12	US-09-814-353-11141
C 14	44	3.8	385	12	US-09-814-353-5368
C 15	44	3.8	385	12	US-09-814-353-11655
C 16	44	3.8	455	10	US-09-728-444-151

17	43.6	3.7	698	12	US-10-029-386-22908	Sequence 22908, A
C 18	43.6	3.7	4321	14	US-10-205-823-35	Sequence 35, Appl
C 19	43.6	3.7	4393	14	US-10-198-846-10819	Sequence 10819, A
C 20	43.4	3.7	384	10	US-09-738-626-544	Sequence 544, App
C 21	43.4	3.7	3302	14	US-10-037-270-475	Sequence 475, App
C 22	43.4	3.7	3309400	10	US-09-738-626-1	Sequence 1, Appli
C 23	43.2	3.7	293	9	US-09-864-761-18923	Sequence 18923, A
C 24	43.2	3.7	459	9	US-09-864-761-2182	Sequence 2182, Ap
C 25	43.2	3.7	1791	9	US-09-416-384A-6	Sequence 6, Appli
C 26	42.8	3.7	586	12	US-10-029-386-9094	Sequence 9094, Ap
C 27	42.8	3.7	4248	14	US-10-198-846-11546	Sequence 11546, A
C 28	42.8	3.7	3673778	12	US-10-312-841-1	Sequence 1, Appli
C 29	42.6	3.6	388	12	US-09-814-353-17525	Sequence 17525, A
C 30	42.6	3.6	797	12	US-09-814-353-4971	Sequence 4971, Ap
C 31	42.6	3.6	797	12	US-09-814-353-11264	Sequence 11264, A
C 32	42.6	3.6	2424	13	US-10-029-217A-25	Sequence 25, Appl
C 33	42.6	3.6	4959	13	US-10-029-217A-1	Sequence 1, Appli
C 34	42.6	3.6	4960	13	US-10-029-217A-30	Sequence 30, Appl
C 35	42.4	3.6	1925	14	US-10-219-420-217	Sequence 217, App
C 36	42.4	3.6	2453	12	US-10-241-009-22	Sequence 22, Appl
C 37	42.4	3.6	2453	12	US-10-190-434B-22	Sequence 237, App
C 38	42.2	3.6	5257	12	US-10-007-926A-237	Sequence 1996, Ap
C 39	42.2	3.6	5959	10	US-09-954-456-1996	Sequence 1996, Ap
C 40	42	3.6	503	12	US-09-814-353-5281	Sequence 5281, Ap
C 41	42	3.6	503	12	US-09-814-353-11568	Sequence 11568, A
C 42	42	3.6	539	12	US-09-814-353-17862	Sequence 17862, A
C 43	42	3.6	632	12	US-09-814-353-5078	Sequence 5078, Ap
C 44	42	3.6	632	12	US-09-814-353-11370	Sequence 11370, A
C 45	41.8	3.6	3673778	12	US-10-312-841-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-918-995-3342  
; Sequence 3342, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3342  
; LENGTH: 474  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(474)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-3342

Query Match	23.0%	Score	269;	DB	11;	Length	474;
Best Local Similarity	98.2%;	Pred. No.	1.9e-70;				
Matches	272;	Conservative	0;	Mismatches	5;	Indels	0;
Gaps	0;						
QY	892	ATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAAA	951				
Db	54	ACTTTTAAGGAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAAA	113				
QY	952	TGAAAACCTCCGGCTACAAATTCAGATGACAAAGACACACACCGCTTCTCTCCCTCC	1011				
Db	114	TGAAAACCTCCGGCTACAAATTCAGATGACAAAGACACACACCGCTTCTCTCCCTCC	173				
QY	1012	CAAAATCCTGTAATCCAGTGATCCTGACAAAGAAAGACAGTAGCTGTGGCCCTCTCACA	1071				



```
; APPLICANT: Pfizer Inc.
; APPLICANT: Castleberry, Tessa A.
; APPLICANT: Lu, Bihong
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: The Canine Androgen Receptor
; FILE REFERENCE: PC10893AGPR
; CURRENT APPLICATION NUMBER: US/10/008,739A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3577
; TYPE: DNA
; ORGANISM: Canine
; US-10-008-739A-1

Query Match      3.9%; Score 46; DB 13; Length 3577;
Best Local Similarity 65.7%; Pred. No. 0.016;
Matches 67; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 769 GCTTCTGACCGCAGCTGTTACTTTACAATACGCTGCTGCTGCGCGCTGTTGTTG 828
      |||||
Db 633 GCTGCTGCTACCTTCTGATACCTCTGCTGTTGCTGCTGCTGCTGCTGTTGCTG 574

QY 829 TGCTGCACTGCTGCTGCGCTGTTGTTTCTGCTGTAGAAG 870
      |||||
Db 573 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAAGAG 532

RESULT 5
US-09-849-243-13/c
; Sequence 13, Application US/09849243
; Patent No. US20020157127A1
; GENERAL INFORMATION:
; APPLICANT: Kirschbaum, Bernd
; Berglund, Erick
; Meisterernst, Michael
; Polites, Greg
; TITLE OF INVENTION: COMPLEXES FROM TRANSGENIC
; NON-HUMAN ANIMALS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HELLER, EHRMAN, WHITE & McAULIFFE
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/849,243
; FILING DATE: 07-May-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Granados, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 38005-0148
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)912-2000
; TELEFAX: (202)912-2020
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
```

```
; NAME/KEY: exon
; LOCATION: 1..1310
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-849-243-13

Query Match      3.8%; Score 44.8; DB 10; Length 1310;
Best Local Similarity 81.2%; Pred. No. 0.019;
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGCTGTTGTTGGTGGCAACTGCTGCTGCGTTGTTTCT 860
      |||||
Db 336 GCTGCTGCTGCTGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGTTGCT 277

QY 861 GCTG 864
      ||||
Db 276 GCTG 273

RESULT 6
US-09-822-846-491/c
; Sequence 491, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 491
; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-822-846-491

Query Match      3.8%; Score 44.8; DB 11; Length 2614;
Best Local Similarity 76.4%; Pred. No. 0.029;
Matches 55; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGCTGTTGTTGGTGGCAACTGCTGCTGCGTTGTTTCT 860
      |||||
Db 2179 GCTGTTGCTGCTGCTGTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGTTGCT 2120

QY 861 GCTGTAGAAGAA 872
      ||||
Db 2119 GCTGCTGCCGCA 2108

RESULT 7
US-09-849-243-15/c
; Sequence 15, Application US/09849243
; Patent No. US20020157127A1
```





QY 913 AGAGAGACAAACAAGAACTGAGACAGAAAGTGGAAATGAAACTCCGGCTACAATTC 972  
Db 129 AA 70  
QY 973 AGATGACAAAGACACACACACA 995  
Db 69 AAAAAAAAAAAAAAAAAAGAGAAAAA 47

RESULT 13  
US-09-814-353-11141/c  
; Sequence 11141, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 11141  
; LENGTH: 299  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 146, 163, 175, 190, 191, 192, 195, 198, 199, 201, 202, 203,  
; LOCATION: 206, 213, 214, 217, 222, 224, 225, 226, 229, 232, 233, 239,  
; LOCATION: 246, 247, 249, 254, 255, 256, 264, 268, 274, 284, 286  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-11141

Query Match 3.8%; Score 44; DB 12; Length 299;  
Best Local Similarity 55.9%; Pred. No. 0.013;  
Matches 80; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 853 TTGTTTCTGCTGTAGAGAAAAAGAGGATTTTCGTATTCAATTTCAAAAGAAATCTGAAA 912  
Db 189 TTTTCTTTTNTTNAAGAAAAAGGGTNTTTTNTTTTNTTNAAGAAAAAGAAAAA 130  
QY 913 AGAGAGACAAACAAGAACTGAGACAGAAAGTGGAAATGAAACTCCGGCTACAATTC 972  
Db 129 AA 70  
QY 973 AGATGACAAAGACACACACACA 995  
Db 69 AAAAAAAAAAAAAAAAAAGAGAAAAA 47

RESULT 14  
US-09-814-353-5368/c  
; Sequence 5368, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela

; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 5368  
; LENGTH: 385  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 148, 149, 160, 161, 162, 163, 173, 174, 175, 176, 177, 178,  
; LOCATION: 179, 181, 192, 193, 210, 213, 214, 215, 216, 218, 219, 234,  
; LOCATION: 238, 239, 240, 241, 242, 245, 247, 254, 256, 261, 262, 264,  
; LOCATION: 265, 267, 268, 272, 273, 274, 275, 276, 277, 278, 279  
; OTHER INFORMATION: n = A,T,C or G  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 280, 281, 282, 283, 284, 285, 288, 289, 301, 303, 304, 305,  
; LOCATION: 306, 307, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,  
; LOCATION: 319, 320, 321, 322, 326, 328, 329, 341, 354, 357, 358, 361,  
; LOCATION: 362, 363, 367, 369, 371, 372, 373, 374, 375, 377, 382  
; OTHER INFORMATION: n = A,T,C or G  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 384  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-5368  
Query Match 3.8%; Score 44; DB 12; Length 385;  
Best Local Similarity 52.7%; Pred. No. 0.015;  
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 850 TTGTTTCTGCTGTAGAGAAAAAGAGGATTTTCGTATTCAATTTCAAAAGAAATCTGA 909  
Db 189 TTTTCTGTTTNN 130  
QY 910 AAGAGAGACAAACAAGAACTGAGACAGAAAGTGGAAATGAAACTCCGGCTACA 969  
Db 129 AA 70  
QY 970 TCCAGATGACAAAGACACACACA 995  
Db 69 AAAAAAAAAAAAAAAAAAGAGAAAAA 44

RESULT 15  
US-09-814-353-11655/c  
; Sequence 11655, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

```
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11655
LENGTH: 385
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 148, 149, 160, 161, 162, 163, 173, 174, 175, 176, 177, 178,
LOCATION: 179, 181, 192, 193, 210, 213, 214, 215, 216, 218, 219, 234,
LOCATION: 238, 239, 240, 241, 242, 245, 247, 254, 256, 261, 262, 264,
LOCATION: 265, 267, 268, 272, 273, 274, 275, 276, 277, 278, 279
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: 280, 281, 282, 283, 284, 285, 288, 289, 301, 303, 304, 305,
LOCATION: 306, 307, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,
LOCATION: 319, 320, 321, 322, 326, 328, 329, 341, 354, 357, 358, 361,
LOCATION: 362, 363, 367, 369, 371, 372, 373, 374, 375, 377, 382
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: 384
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-11655

Query Match      3.8%; Score 44; DB 12; Length 385;
Best Local Similarity 52.7%; Pred.No. 0.015;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      850 TTGTTGTTCTGCTGTAGAGAAAGACGAGGTTTCGTTATTCAATTTCAAAGAAATCTGA 909
Db      189 TTTTGTGTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 130

QY      910 AAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAATGAAATCAAACTCCGGCTACAA 969
Db      129 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 70

QY      970 TTCAGATGACAAAGACACACACA 995
Db      69 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 44
```

Search completed: November 13, 2003, 03:28:16  
Job time : 397.507 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 02:52:59 ; Search time 34.1179 Seconds  
(without alignments)  
2919.536 Million cell updates/sec

Title: US-09-729-264-4

Perfect score: 2088

Sequence: 1 MVAGAMENRPPGSGSNEV.....HPQAFNLASPEKVSNTTV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*\*  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_rvirus:  
16: sp\_bacteriap:  
17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1692	81.0	315	4 Q9NS15	Q9ns15 homo sapien
2	566.5	27.1	11	Q9D8G2	Q9d8g2 mus musculus
3	147	7.0	492	11 Q8C0U8	Q8c0u8 mus musculus
4	141	6.8	168	11 Q9D9I2	Q9d9i2 mus musculus
5	138	6.6	168	11 Q8CH20	Q8ch20 mus musculus
6	138	6.6	1496	4 Q92626	Q92626 homo sapien
7	136.5	6.5	173	11 Q9D4K2	Q9d4k2 mus musculus
8	135	6.5	1445	11 Q63155	Q63155 rattus norv
9	134.5	6.4	1427	13 Q91562	Q91562 xenopus lae
10	133.5	6.4	871	11 Q61987	Q61987 mus musculus
11	133.5	6.4	881	11 Q61988	Q61988 mus musculus
12	132.5	6.3	869	4 Q15146	Q15146 homo sapien
13	129.5	6.2	164	5 Q22048	Q22048 caenorhabdi
14	127	6.1	197	5 Q17641	Q17641 caenorhabdi
15	125.5	6.0	166	5 Q95QV1	Q95qv1 caenorhabdi
16	125	6.0	344	13 Q9DF61	Q9df61 gallus gall

#### ALIGNMENTS

RESULT 1

Q9NS15 Q9NS15 PRELIMINARY; PRT; 315 AA.

AC Q9NS15;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
-T 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE IGSF5 protein (Fragment).  
GN IGSF5.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

SEQUENCE FROM N.A.  
RP Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,  
RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,  
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,  
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
RA Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,  
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,  
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,  
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,  
RA Ramser J., Beck A., Klages S., Hennig S., Rieseemann L., Degand E.,  
RA Wehrmeyer S., Borzym K., Gardiner K., Yaspo M.L.;  
RA Leinhardt H., Reinhardt R., Yaspo M.L.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
EMBL; AL163280; CAB90447.1; -  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR003599; Ig\_like.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00409; IG; 2.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS00835; IG\_LIKE; 2.  
FT NON\_TER 1 315

Q9Z290 mus musculu  
Q9dbp0 mus musculu  
Q18238 caenorhabdi  
Q8nc34 homo sapien  
Q8nb18 homo sapien  
Q9eaa2 homo sapien  
Q9u1r0 homo sapien  
Q8bj62 mus musculu  
Q57596 gallus gall  
Q9dgi5 gallus gall  
Q93242 gallus gall  
Q9w213 drosophila  
Q44924 drosophila  
Q8ji27 brachydanio  
Q8ay67 brachydanio  
Q01761 caenorhabdi  
Q8nf45 homo sapien  
Q8nf46 homo sapien  
Q8nf48 homo sapien  
Q96rw7 homo sapien  
Q62838 rattus norv  
Q8nda2 homo sapien  
Q9vw64 drosophila  
Q8msq0 drosophila  
Q8axy6 gallus gall  
Q8msn7 drosophila  
Q9w4t9 drosophila  
Q9n9y9 drosophila  
Q97174 drosophila











OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96009854; PubMed=7546737;  
 RA Valenzuela D.M., Stitt T.N., DiStefano P.S., Rojas E., Matteson K.,  
 RA Compton D.L., Nunez L., Park J.S., Stark J.L., Gies D.R., Thomas S.,  
 RA LeBeau M.M., Fernald A.A., Copeland N.G., Jenkins N.A., Burden S.J.,  
 RA Glass D.J., Yancopoulos G.D.;  
 RT "Receptor tyrosine kinase specific for the skeletal muscle lineage:  
 RT expression in embryonic muscle, at the neuromuscular junction, and  
 RT after injury.";  
 RL Neuron 15:573-584 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Valenzuela D.M., Rojas E., Yancopoulos G.D.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF006464; AAB63044.1; -;  
 DR HSP; P11362; 1FGK.  
 DR Genes; HGNC:7525; MUSK.  
 DR InterPro; IPR002453; Beta\_tubulin.  
 DR InterPro; IPR000024; P2\_domain.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003598; IG\_c2.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF01392; Pz; 1.  
 DR Pfam; PF00047; Ig; 3.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR0109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00408; IGC2; 3.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00238; Pz; 1.  
 DR PROSITE; PS00835; IG-LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; 1.  
 KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase;  
 KW Tyrosine-protein kinase  
 SQ SEQUENCE 869 AA; 97056 MW; 3DDC20E179FA010C CRC64;  
 Query Match 6.3%; Score 132.5; DB 4; Length 869;  
 Best Local Similarity 20.8%; Pred. No. 0.0018;  
 Matches 89; Conservative 55; Mismatches 178; Indels 105; Gaps 20;  
 QY 4 GAGNEDPPGSGNEVIEGPQNAVRLKGSQARENCTVSGQWK-LIMWALSMDVLSVRP 62  
 DB 113 GALQVKKPK-----KITPPINVKIIEGLKAVLPCTTNGNPKPSVSWIKGD----- 158  
 QY 63 MEPIITNDRTSORYDQGNFTSEMIHNVPSPDSGNIRCSLQNGRLHGSAY-LTVQVWG 121  
 DB 159 -SPLRENSRIAVLE-----SSSLRIHNQVEDAGQYRCVAKNSL--GTATSKVVKLEV 208  
 QY 122 ELFPISVNLVVAENP-----CEVTCLPSHTWLPDISW-ELGLVSHSHSYFVPEP 172  
 DB 209 EVFARILRAPESHNVITGFGVTLHCTATGIP-----VPTITWENGNVSSGSIQESVKD 263  
 QY 173 SDIQSAVSIALLTPQSGNTLTCTVAT-----WKSILKARKSATVNLTVIRCPQDTGGG----- 223  
 DB 264 RVIDSRILQLPITKP--GLYTCIATNKGKFKSTAKAATISIAEWSKPKQKNGVCAQY 320  
 QY 224 -----INIPGVLSLPSLGSFL-----PTWKGVLGLAGTMLTPTCTLTIRCC 267  
 DB 321 RGEVCNAVLAADALVFLNTSYADPEBAQELLVHTAWNEL-----KVVSPPVCRFAEAL 373  
 QY 268 CRRRCRCGN-----CCRCRC-----FCRRKRGRIQFOKKSEKENTKETETE 312  
 DB 374 LCNHIFQECSPGVVPTPIPCREYCLAVKELFCAKE-----WLWEEKTHRGLYRSEMHLL 429

QY 313 SGNENSGYNSDEQKTTDTASLPKSCSSDPEQRNSSCGPPHQADQRP-----PRPASHP 368  
 DB 430 SVPECKSLPSMHWDPACARLP-----HLDYKNENLKTPTP--MTSSKFSVDIPNLPSSS 482  
 QY 369 QASFNLA 375  
 DB 483 SSSFSVS 489  
 RESULT 13  
 Q22048 PRELIMINARY; PRT; 164 AA.  
 AC Q22048;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE TUB7.8 protein.  
 GN TUB7.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sims M.A.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 DR EMBL; Z66499; CAA91301.1; -;  
 DR WormPep; T01B7.8; CE03592.  
 DR InterPro; IPR001450; 4Fe4S\_ferredoxin.  
 DR InterPro; IPR005081; Defensin\_alpha.  
 DR InterPro; IPR006209; EGF-like.  
 DR PROSITE; PS00198; 4Fe4S\_FERREDOXIN; 1.  
 DR PROSITE; PS00269; DEFENSIN; 1.  
 DR PROSITE; PS00222; EGF\_1; 1.  
 SQ SEQUENCE 164 AA; 16499 MW; C002D48D36C9FCBD CRC64;  
 Query Match 5.2%; Score 129.5; DB 5; Length 164;  
 Best Local Similarity 32.1%; Pred. No. 0.00039;  
 Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;  
 QY 157 LGLVSHSHSYFVPEPDLQSAVSIALLTPQSGNTLTCTVATKSKARKSATVNLTVIRC 216  
 DB 6 LAILLAIGTPTAV--SQVQSAV-----LPVSSTELATVGTVDVSTASTAIDTLGNSSRV 57  
 QY 217 PDGTGGGINIPGVLSLPSLGSFLPTWKGVLGLAGTMLTPTCTLTIRCCRRRCGC 276  
 DB 58 KRQGGCGCCGCGC-----GCCCGCGGGGG--CGCCCRPRCCCCRCCTC 101  
 QY 277 --NCCC-RCCFCR 287  
 DB 102 CRTCCCTRCCTCR 115  
 RESULT 14  
 Q17641 PRELIMINARY; PRT; 197 AA.  
 AC Q17641;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical 20.6 kDa protein.  
 GN C04G6.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;

```

Query Match          6.0%; Score 125.5; DB 5; Length 166;
Best Local Similarity 64.5%; Pred. No. 0.00095;
Matches 20; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy      260 CTUTIRCCCRRRCCGC--NCCC-RCFCOCR 287
      | | | | | | | | | | | | | | | |
Db      86 CCCRPRCCCCCRCCCTCCCTCCCTCCCTCCR 116

Search completed: November 13, 2003, 03:33:36
Job time : 36.4512 secs

```

```

RESULT 15
Q95QY1
ID Q95QY1 PRELIMINARY; PRT; 166 AA.
AC Q95QY1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 17.0 kDa protein.
C04G6.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium." ;
RL Science 282:2012-2018(1998) .
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Anderson K., Chisoe S.;
RT "The sequence of C. elegans cosmid C04G6." ;
RL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 23:19:44 ; Search time 34.1179 Seconds  
(without alignments)  
1795.789 Million cell updates/sec

Title: US-09-729-264-4  
Perfect score: 2088  
Sequence: 1 MVAGAMENRPPGSGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158756573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2088	100.0	386	AAU75541	Human B7-like prot
2	2004	96.0	382	AAU75540	Human B7-like prot
3	2000	95.8	386	AAU75542	Human B7-like prot
4	1524	75.5	377	AAU75543	Human B7-like prot
5	1318	63.1	463	ABG28169	Novel human diagno
6	913.5	43.8	370	AAU75544	Mouse B7-like prot
7	572.5	27.4	631	AAU75547	Rat B7-like prot
8	566.5	27.1	270	AAU75545	Mouse B7-like prot
9	485	23.2	223	AAU75546	Mouse B7-like prot

10	157	7.5	420	23	ABP65012	Human protein SEQ
11	145.5	7.0	404	22	ABP65012	Extracellular c1r
12	145.5	7.0	404	23	ABP65011	Human protein SEQ
13	145.5	7.0	404	23	AAE23219	Human receptor for
14	145.5	7.0	404	23	AAU77543	Human receptor for
15	145.5	7.0	404	23	AAU48745	Human RAGE protein
16	138	6.6	1496	20	AAU81030	Melanoma associat
17	138	6.6	1496	21	AAU70469	Human p53 target m
18	138	6.6	1496	22	ABU03498	Angiogenesis-asso
19	138	6.6	1498	22	ABU11587	Human peroxidase
20	136	6.5	1447	16	AAE68553	Deleted in colorec
21	136	6.5	1447	20	AAU33496	Human DCC protein.
22	136	6.5	1447	22	AAU50693	Human UNC-40 prote
23	136	6.5	1447	24	ABJ19765	Human MP21 protein
24	136	6.5	1728	12	AAU13144	Deleted in Colorec
25	133.5	6.4	467	19	AAU62575	Alternatively spli
26	133.5	6.4	475	17	AAU94982	Nsk2 extracellular
27	133.5	6.4	863	19	AAU62569	Alternatively spli
28	133.5	6.4	867	19	AAU62583	Mouse receptor tyr
29	133.5	6.4	871	17	AAU84087	Nsk2 receptor. Mu
30	133.5	6.4	873	17	AAU84092	Mouse receptor tyr
31	133.5	6.4	873	19	AAU62573	Nsk2 receptor with
32	133.5	6.4	881	17	AAU84091	Alternatively spli
33	133.5	6.4	881	19	AAU62572	Mouse Nsk2 (altern
34	133.5	6.3	869	18	AAU26611	Human muscle-speci
35	132.5	6.3	869	18	AAU26506	Human Dmk receptor
36	132.5	6.3	633	24	AAU26267	gp354 clone protei
37	132	6.3	633	22	AAU77857	Mutant protein mmu
38	131.5	6.3	537	22	ABG74129	Human hMusK-R dele
39	131.5	6.3	537	24	ABG74129	Mutant protein mmu
40	131.5	6.3	576	22	ABG77858	Human hMusK-R dele
41	131.5	6.3	576	24	ABG74130	Protein of muscle
42	131.5	6.3	869	22	AAU77856	Amino acid sequenc
43	131.5	6.3	869	22	AAU68421	Human muscle speci
44	131.5	6.3	869	24	ABG74128	Novel human diagno
45	130.5	6.2	1483	22	ABG16336	

## ALIGNMENTS

RESULT 1  
AAU75541  
ID AAU75541 standard; Protein; 386 AA.  
XX AAU75541;  
XX AC  
XX DT 23-APR-2002 (first entry)  
XX DE Human B7-like protein, B7-L\_h2.  
XX KW Human; B7-like protein; B7-L; antiinfertility; gynaecological;  
XX KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
XX KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
XX KW antidiabetic; haemostatic; antichyroid; antiulcer; antiallergic;  
XX KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
XX KW reproductive disorder; graft versus host disease; autoimmune disease;  
XX KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
XX KW endocrinopathy; lymphoproliferative disorder.  
XX OS Homo sapiens.  
XX PN WO200200710-A2.  
XX PD 03-JAN-2002.  
XX PF 28-JUN-2001; 2001WO-US20719.  
XX PR 28-JUN-2000; 2000US-214512P.  
XX PR 28-NOV-2000; 2000US-0729264.  
XX PA (AMGE-) AMGEN INC.

Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
WPI; 2002-130881/17.  
N-PSDB; ABK13029.  
New B7-like polypeptides, polynucleotides and their modulators, useful  
for diagnosing, preventing and treating reproductive, immune and  
proliferative disorders, e.g. cancer and arteriosclerosis -  
Claim 13; Fig 2; 135pp; English.  
The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
The polypeptide, polynucleotide encoding it and antibody against (I) are  
useful for treating B7-like polypeptide-related disease, disorders or  
conditions including reproductive disorders (e.g. infertility, and  
miscarriage, preterm labour and delivery and endometriosis) and  
proliferative disorders. Antibodies, soluble proteins comprising  
extracellular domains and other regulators of B7-L polypeptides are  
useful for enhancing the immune response to tumours. (I) plays a role in  
growth and maintenance of cancer cells based on the observation of  
seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
polypeptide. Hence modulators of (I) are useful for the treatment of  
cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
in allograft transplantation, graft versus host disease, T-cell  
dependent B-cell mediated diseases and autoimmune diseases. B7-L  
molecules are useful for alleviating the symptoms associated with  
diseases involving chronic immune cell dysfunction or to treat  
autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
purpura and psoriasis, chronic inflammatory disease such as  
inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
are also useful as immunosuppressive agents for bone marrow and organ  
transplantation or to prolong graft survival. B7-L molecules are also  
useful for diagnosis and treatment of diseases involving abnormal cell  
proliferation, including arteriosclerosis and vascular restenosis.  
Antagonists of B7-L polypeptides are useful for alleviation of toxic  
shock syndrome or allo sensitization due to blood transfusions, and for  
treatment of allergy, asthma and hypersensitivity reactions,  
nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
(extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and  
thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
lymphoproliferative disorders such as multiple myeloma. The present  
sequence represents the amino acid sequence of human B7-L<sub>h2</sub>.  
Sequence 386 AA;  
Query Match 100.0%; Score 2088; DB 23; Length 386;  
Best Local Similarity 100.0%; Pred. No. 6.8e-165;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVAGAMENRDPGSGNGNEVIEGPQARVLKSGQARFNCTVSQGWKLIMWALSVMVLSV 60  
DB 1 MVAGAMENRDPGSGNGNEVIEGPQARVLKSGQARFNCTVSQGWKLIMWALSVMVLSV 60  
QY 61 RMPEIITNDRTSQRYDQGNFTSEMIHNHVPDSGNIRCSLQNSRLHGSAYLTVQVM 120  
DB 61 RMPEIITNDRTSQRYDQGNFTSEMIHNHVPDSGNIRCSLQNSRLHGSAYLTVQVM 120  
QY 121 GELFIPSNVLVAENPEVTCPLPSHTWLPDISWELGLLVSHSGSYFVPPSPDLQAVS 180  
DB 121 GELFIPSNVLVAENPEVTCPLPSHTWLPDISWELGLLVSHSGSYFVPPSPDLQAVS 180  
QY 181 ILALTPQSGNLTCTVATWKSLSKARKSATVNLTRCPQDTGGGINPQVLSLPSLGFSL 240  
DB 181 ILALTPQSGNLTCTVATWKSLSKARKSATVNLTRCPQDTGGGINPQVLSLPSLGFSL 240  
QY 241 PTWKGVLGLAGTMTLLTPCTLTITICCCRRCCGNCNCCRCCKRKRFRIFQOKKS 300  
DB 241 PTWKGVLGLAGTMTLLTPCTLTITICCCRRCCGNCNCCRCCKRKRFRIFQOKKS 300

QY 301 EKEKTKETETESGNGNSGYNSDEKTTDTASLPKSCSSDPQRNSSCGPPHORADOR 360  
DB 301 EKEKTKETETESGNGNSGYNSDEKTTDTASLPKSCSSDPQRNSSCGPPHORADOR 360  
QY 361 PPRPASHPQASFNLASPEKVSNTTVV 386  
DB 361 PPRPASHPQASFNLASPEKVSNTTVV 386  
RESULT 2  
AAU75540  
ID AAU75540 standard; Protein; 382 AA.  
XX  
AC AAU75540;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE Human B7-like protein, B7-L<sub>h1</sub>.  
XX  
KW Human; B7-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder.  
XX  
OS Homo sapiens.  
XX  
XX WO200200710-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 28-JUN-2001; 2001WO-US20719.  
XX  
PR 28-JUN-2000; 2000US-214512P.  
PR 28-NOV-2000; 2000US-0729264.  
XX  
XX (AMGE-) AMGEN INC.  
XX  
PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
XX  
XX WPI; 2002-130881/17.  
DR N-PSDB; ABK13028.  
XX  
XX New B7-like polypeptides, polynucleotides and their modulators, useful  
PT for diagnosing, preventing and treating reproductive, immune and  
PT proliferative disorders, e.g. cancer and arteriosclerosis -  
XX  
PS Claim 13; Fig 1; 135pp; English.  
XX  
CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
CC useful for treating B7-like polypeptide-related disease, disorders or  
CC conditions including reproductive disorders (e.g. infertility, and  
CC miscarriage, preterm labour and delivery and endometriosis) and  
CC proliferative disorders. Antibodies, soluble proteins comprising  
CC extracellular domains and other regulators of B7-L polypeptides are  
CC useful for enhancing the immune response to tumours. (I) plays a role in  
CC growth and maintenance of cancer cells based on the observation of  
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
CC polypeptide. Hence modulators of (I) are useful for the treatment of  
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
CC in allograft transplantation, graft versus host disease, T-cell  
CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
CC molecules are useful for alleviating the symptoms associated with  
CC diseases involving chronic immune cell dysfunction or to treat  
CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic



QY 133 AENPCEVTLCPSHWTWLPDISWELGLLVSHSSYFVPEPSDLQSAVSILALTPQSNGLT 192  
 Db 133 AENPCEVTLCPSHWTWLPDISWELGLLVSHSSYFVPEPSDLQSAVSILALTPQSNGLT 192  
 QY 193 TCVAATWSLKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGFSLPTWGWKVLGLAG 252  
 Db 193 TCVAATWSLKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGFSLPTWGWKVLGLAG 252  
 QY 253 TMLLTPTCTLTIRCCCRRCGCCGCCRCFCRRRKGFRIFQKSEKEKTKNKETE 312  
 Db 253 TMLLTPTCTLTIRCCCRRCGCCGCCRCFCRRRKGFRIFQKSEKEKTKNKETE 312  
 QY 313 SGHNSGYNDEQKTTTASLPKPKSCSSDPQQRNSCGPPHQRADQRRPPASHPOASF 372  
 Db 313 SGHNSGYNDEQKTTTASLPKPKSCSSDPQQRNSCGPPHQRADQRRPPASHPOASF 372  
 QY 373 NLASPEKVSNTTVV 386  
 Db 373 NLASPEKVSNTTVV 386

RESULT 4  
 AAU75543  
 ID AAU75543 standard; Protein; 377 AA.  
 AC AAU75543;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE Human B7-L-like protein, B7-L\_h4.  
 XX  
 KW Human; B7-L; antiinfertility; gynaecological;  
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;  
 KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200710-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 28-JUN-2001; 2001WO-US20719.  
 XX  
 PR 28-JUN-2000; 2000US-214512P.  
 PR 28-NOV-2000; 2000US-0729264.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
 XX  
 DR WPI; 2002-130881/17.  
 DR N-PSDB; ABK13031.  
 XX  
 PT New B7-like polypeptides, polynucleotides and their modulators, useful  
 PT for diagnosing, preventing and treating reproductive, immune and  
 PT proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX  
 PS Claim 13; Fig 4; 135pp; English.  
 XX  
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
 CC useful for treating B7-like polypeptide-related disease, disorders or  
 CC conditions including reproductive disorders (e.g. infertility,  
 CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (I) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of

CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions, and for  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the amino acid sequence of human B7-L\_h4.  
 XX

SQ Sequence 377 AA;

Query Match 75.5%; Score 1576; DB 23; Length 377;  
 Best Local Similarity 88.6%; Pred. No. 2.1e-122;  
 Matches 296; Conservative 12; Mismatches 20; Indels 6; Gaps 2;  
 QY 1 MVAGAMENRDPGSGSGNEVIEGPQARVNLKGSQARFNCTVSGWKLIMWALSDMVVLSV 60  
 Db 1 MVAGAMENRDPGSGSGNEVIEGPQARVNLKGSQARFNCTVSGWKLIMWALSDMVVLSV 60  
 QY 61 RMEPIITNDRFTSQRYDOGNGFTSEMIITHNVEPSDSGNIRCSLQNSRLHGSAYLTQVM 120  
 Db 61 RMEPIITNDRFTSQRYDOGNGFTSEMIITHNVEPSDSGNIRCSLQNSRLHGSAYLTQVM 120  
 QY 121 GELFIPSVNLVVAENEPCEVTCLPSHWTWLPDISWELGLLVSHSSYFVPEPSDLQSAVS 180  
 Db 121 GELFIPSVNLVVAENEPCEVTCLPSHWTWLPDISWELGLLVSHSSYFVPEPSDLQSAVS 180  
 QY 181 ILALTPQSNGLTTCVATWKSILKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGFSL 240  
 Db 181 ILALTPQSNGLTTCVATWKSILKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGFSL 240  
 QY 241 PTWGWKVLGLAGTMLLTPTCTLTIRCCCRRCGCCGCCRCFCRRRKGFRIFQFQKS 300  
 Db 241 PTWGWKVLGLAGTMLLTPTCTLTIRCCCRRCGCCGCCRCFCRRRKGFRIFQFQKS 300  
 QY 301 EKEKTKETETESGNSGYNDEQKTTTASLP 334  
 Db 298 QTKLRLQKVEKMT---PATIQMKRPQTPLLSLP 328

RESULT 5  
 ABG28169  
 ID ABG28169 standard; Protein; 463 AA.  
 XX  
 AC ABG28169;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #28160.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.

XX WO200175067-A2.  
 PN 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS92356.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX Claim 20; SEQ ID No 58528; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 463 AA;  
 Query Match 63.1%; Score 1318; DB 22; Length 463;  
 Best Local Similarity 99.6%; Pred. No. 6.9e-101;  
 Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 49 MWALSDMVVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVFSPSGNIRCSLQNSR 108  
 Db 1 MWALSDMVVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVFSPSGNIRCSLQNSR 60  
 QY 109 LHGSAYLTQVMGELFIPSNLVVAENPECVTCPLPSHWTLVDLSWELGLVSHSYYP 168  
 Db 61 LHGSAYLTQVMGELFIPSNLVVAENPECVTCPLPSHWTLVDLSWELGLVSHSYYP 120  
 QY 169 VPEPSDLSQAVSILALTPQSGNLTLCVATWKSLSKARKSATVNLTVIRCPQDTGGGINIPG 228  
 Db 121 VPEPSDLSQAVSILALTPQSGNLTLCVATWKSLSKARKSATVNLTVIRCPQDTGGGINIPG 180  
 QY 229 VLSSLSLGSFLPTWKGVLGLAGTWMLLPTCTLTIRCCRRCCGNCRCFCRCR 288  
 Db 181 VLSSLSLGSFLPTWKGVLGLAGTWMLLPTCTLTIRCCRRCCGNCRCFCRCR 240  
 QY 289 KRQ 291  
 Db 241 KRQ 243

RESULT 6  
 AAU75544  
 ID AAU75544 standard; Protein; 370 AA.  
 XX AC AAU75544;  
 XX DT 23-APR-2002 (first entry)  
 XX DE Mouse B7-like protein, B7-L\_m1.  
 XX KW Mouse; B7-like protein; B7-L; antiinfertility; gynaecological;  
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;  
 KW antiasrhmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder.  
 XX OS Mus musculus.  
 XX PN WO200200710-A2.  
 XX FD 03-JAN-2002.  
 XX PF 28-JUN-2001; 2001WO-US20719.  
 XX PR 28-JUN-2000; 2000US-214512P.  
 XX PR 28-NOV-2000; 2000US-0729264.  
 XX (AMGR-) AMGEN INC.  
 XX Weiher AA, Sarmiento UM, Schultz HJ, Chute HT;  
 WPI: 2002-130881/17.  
 N-PSDB; ABK13032.  
 XX New B7-like polypeptides, polynucleotides and their modulators, useful  
 PT for diagnosing, preventing and treating reproductive, immune and  
 PT proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX Claim 13; Fig 5; 135pp; English.  
 XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
 CC useful for treating B7-like polypeptide-related disease, disorders or  
 CC conditions including reproductive disorders (e.g. infertility,  
 CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (I) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,

CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the amino acid sequence of mouse B7-L<sub>mi</sub>.  
 XX  
 SQ Sequence 370 AA;  
 Query Match 43.6%; Score 913.5; DB 23; Length 370;  
 Best Local Similarity 50.7%; Pred. No. 1.9e-67;  
 Matches 192; Conservative 50; Mismatches 102; Indels 35; Gaps 5;  
 QY 14 SSGNEVIEGPNVAVLQSGARFNCTVSGQKLMWALSMDVLSVRPMEPIITNDRE 73  
 DB 21 SSSYQIIEGPNVAVLQSGARFNCTVSGQKLMWALSMDVLSVRPMEPIITNDRE 80  
 QY 74 SORYDQGNFTSEMIHNVPSDSGNIRCSLQNSRLHGSAYLTVOVMGELFIPSVNLVA 133  
 DB 81 YASYNSTDSEFISELIIHDVQPSDGSVQCSLQNSHGFGSAFLSVQVMGTLNIPSNLIVT 140  
 QY 134 ENPEVETCLPSHWTLWDIHWELGLLYSHSYFVFPSPDLQSAVSLIALTPOSGNLT 193  
 DB 141 EGEPNVCTVAVGWTSLPDISWELEVPVSHSYNFFLEFGNPMRVLSVLDLTPLGNGTLT 200  
 QY 194 CVATWKSLSKARSATVNLTVIRCPDQDTGGINIPGVLSLPSLSLPSLTGKVGGLAGT 253  
 DB 201 CVAELKDLQASKSLVNLTVVQPPD-----SIEEGPALFTWAILLAVAFS 248  
 QY 254 MLTPTCTTITRCOCRRCCGNCRCFCRRKRGFRTOQ---KSEKEKTNK--- 307  
 DB 249 LLLIILVLIIFCC-----CASRRKEESTYQNEIRKSAMTNKADP 293  
 QY 308 ETETSGNNGYNSDEQKTTDTASLPKSCSSDPEORNSCGPPHORADQRP RPASH 367  
 DB 294 ETKLSGKNGYSSDEAKAAQTASLPKSAEVSILPEKSSSL--PYOELNKHQPGPATH 351  
 QY 368 PQASFNLASPEKVSNTTV 386  
 DB 352 PRVSPDIASPKVRNVLV 370

RESULT 7

AAU75547 standard; Protein; 631 AA.

AAU75547;

23-APR-2002 (first entry)

Rat B7-like protein, B7-1.

Rat; B7-like protein; B7-L; antiinfectivity; gynaecological;

antitumour; cytostatic; immunosuppressive; antiarthritic; antitheatumatic;

antiinflammatory; dermatologic; antipsoriatic; neuroprotective;

antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;

antidiabetic; nephrotropic; antibacterial; virucide; tumour; cancer;

reproductive disorder; graft versus host disease; autoimmune disease;

toxic shock syndrome; allergy; nephropathy; skin disorder;

endocrinopathy; lymphoproliferative disorder.

Ratus rattus.

WO200200710-A2.

03-JAN-2002.

28-JUN-2001; 2001WO-US20719.

28-JUN-2000; 2000US-214512P.

28-NOV-2000; 2000US-0729264.

(AMGE-) AMGEN INC.

XX PI Welcher AA, Sarmiento UM, Schultz H7, Chute HT;  
 XX WPI; 2002-130881/17.  
 DR  
 PT New B7-like polypeptides, polynucleotides and their modulators, useful  
 PT for diagnosing, preventing and treating reproductive, immune and  
 PT proliferative disorders, e.g. cancer and arteriosclerosis  
 XX Disclosure; Fig 8; 135pp; English.  
 PS The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 PS The polypeptide, polynucleotide encoding it and antibody against (I) are  
 CC useful for treating B7-like polypeptide-related disease, disorders or  
 CC conditions including reproductive disorders (e.g. infertility,  
 CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (i) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions.  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the amino acid sequence of rat B7-1.

SQ Sequence 631 AA;

Query Match 27.4%; Score 572.5; DB 23; Length 631;

Best Local Similarity 26.5%; Pred. No. 7.6e-39;

Matches 168; Conservative 43; Mismatches 109; Indels 315; Gaps 10;

QY 14 SSGNEVIEGPNVAVLQSGARFNCTVSGQKLMWALSMDVLSVRPMEPIITNDRE 25

DB 21 SSSYQIIEGPNVAVLQSGARFNCTVSGQKLMWALSMDVLSVRPMEPIITNDRE 80

QY 26 NARVLKSGARFNCTVSGQKLMWALSMDVLSVRPMEPIITNDRE 72

DB 81 NVTVLKDSAEAFNCTVTHGKLLMTLNQWLVSLTQGPITNNRFEKALLSCDYKPCS 140

QY 73 -----TSQRYDQGNFTSEMIHNVPSDS 97

DB 141 EQSIHRIYQKHDKMVLVSVISGVPEVPFKYKNTTYASYNSTDSEFISELIIHDVQPSDS 200

QY 98 GNIRCSLQNSRLHGSAYL----- 115

DB 201 GSVQCSLQNSHGFGSAFLSVQVYDIANNYSFSLGLILSDRGTYTCTVQVRYEGSVYWKH 260

QY 116 --TVQVMGELFIPSVNLVVAENEPCEVTCLPSHWTLWDIHWELGLLYSHS----- 164

DB 261 LTTVEVMGTLNIPSNLIVTEGEPNCNVCVAVGNTSLPDISWELEVPVSHSLSVRADPFT 320

165 -----SYFVPEPSDLQSAVSLALT 185  
 321 PNTEYGNPSADIKRITCFASGFFPKPRUSWLENRELSYNFLEPGNPMRVLSVLDLT 380  
 186 POSNGTLTCTVATWKSLSKARKSATVNL----- 211  
 381 PLNGTLTCTVAVELKDLQASKSLTNVLGINTTISQDPESELYTTISSQLDFNATYDHFIDCF 440  
 212 -----TVIRCPQDTGGGINPGVLSLPSGLSPFTWGVGLGLAGTMLL----- 256  
 441 IEYGDHVSQNFVTVOVPPD-----SIGEGPALPTWAILLAVAFSLILLI 488  
 257 -----TPTCLTLTRCC 268  
 489 IVLIIITWKPDPDEKQTVPFAMAGDAVKAIIFIAITVIAVIAAIIIFCCC 548  
 269 CRRCCGNCRCRCFCRCRGRFIOFQ-KKSEKENTK-----ETETESNGNSGNSDE 324  
 549 CASR-----REKEESTYQNEIRKSNMRTNKADPPTKLKSGKENYGYSSDE 594  
 325 QKTDTASLPKSCSSDPQRNSSCGPPHQRADO 359  
 595 CITVFKRCFRFRNEASRETNNLYIGPVAAREQ 629

RESULT 8

AAU75545  
 ID AAU75545 standard; Protein; 270 AA.

XX AAU75545;

XX 23-APR-2002 (first entry)

XX Mouse B7-like protein, B7-L m2.

XX Mouse; B7-like protein; B7-L; anti-infectivity; gynaecological;  
 XX antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 XX antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 XX antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
 XX antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
 XX reproductive disorder; graft versus host disease; autoimmune disease;  
 XX toxic shock syndrome; allergy; nephropathy; skin disorder;  
 XX endocrinopathy; lymphoproliferative disorder.

XX Mus musculus.

XX WO200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.

XX 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.

XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX WPI; 2002-130881/17.

XX N-PSDB; ABK13033.

XX New B7-like polypeptides, polynucleotides and their modulators, useful  
 XX for diagnosing, preventing and treating reproductive, immune and  
 XX proliferative disorders, e.g. cancer and arteriosclerosis -

XX Claim 13; Fig 6; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
 XX useful for treating B7-like polypeptide-related disease, disorders or  
 XX conditions including reproductive disorders (e.g. infertility,

CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (I) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various neuropathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the amino acid sequence of mouse B7-L m2.

XX Sequence 270 AA;

Query Match 27.1%; Score 566.5; DB 23; Length 270;

Best Local Similarity 34.6%; Pred. No. 8.2e-39;

Matches 131; Conservative 42; Mismatches 71; Indels 135; Gaps 5;

QY 14 SSGNEVIEGPQARVLRKSGQARFNCVTSQGWKLIMWALSMDVLSVRPMEPIITNDRPT 73

DB 21 SSSYQIIIEGPQVTVLRKSDSEAHFNCTVTRGKLLMWTLLNQWVLSLTATQGPFIITNDRPT 80

QY 74 SORYQGGNFTSEMIHNVPDSDSNIRCSLNSRLHGSAYLTVCVMSELPIPSVNLVVA 133

DB 81 YASYNSTDSFISLIIHDVQPSDSGVCSQNSHGFGSAPLSVQ----- 125

QY 134 ENPECEVTCLPSSHWTWLPDISWELGLLVSHSSYFVPEPFDIQLQSAVSLALTPQSNGLT 193

DB 126 ----- 125

QY 194 CVATWKSLSKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLPSLPTWGVGLGLACT 253

DB 126 -----DSIGEGPALPTWAILLAVAFS 148

QY 254 MLLTPTCTLTIRCCCRRCRCGCCNCCRCFCRCRGRFIOFQ---KKSEKENTK--- 307

DB 149 LLLIILVLIIFCCC-----CASREKEESTYQNEIRKSNMRTNKADP 193

QY 308 ETETESNGNSGNSDEQKTTDTASLPKSCSSDPQRNSSCGPPHQRADORPPRPASH 367

DB 194 ETKLKGKENYGYSSDEAKAAQATASLPKSAEVSILPEKRSSSL--PYQLNKHQFGPATH 251

QY 368 PQASFNLSAPEKVSNTTVV 386

DB 252 PRVSFDIASPQKVRNVTLV 270

RESULT 9

AAU75546

ID AAU75546 standard; Protein; 223 AA.

XX AAU75546;

AC AAU75546;





ABP65011	ABP65011 standard; Protein; 404 AA.	
AC	ABP65011;	
DT	25-FEB-2003 (first entry)	
XX	Human protein SEQ ID 671.	
DE	Human; expressed sequence tag; EST;	
XX	haematopoietic disorder; central nervous system disease; viral infection;	
KW	peripheral nervous system disease; non-healing wound; infectious disease;	
KW	immune deficiency; immune disorder; bacterial infection; allergy; cancer;	
KW	fungal infection; autoimmune disorder; coagulation disorder; neutropenic;	
KW	antiallergic; antiinflammatory; immunosuppressive; neuroprotective;	
KW	cytostatic; haemostatic; virucide; antibacterial; fungicide;	
KW	immunostimulant; cerebroprotective.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200259260-A2.	
XX		
PD	01-AUG-2002.	
XX		
PF	16-NOV-2001; 2001WO-US42950.	
XX		
PR	17-NOV-2000; 2000US-0714936.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;	
PI	Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;	
XX		
DR	WPI; 2002-590824/63.	
DR	N-PSDB; ABQ95937.	
XX		
PT	New isolated polynucleotide, useful in research, diagnostic or	
PT	therapeutic methods, e.g. preventing or treating disorders involving	
PT	aberrant protein expression or biological activity -	
XX		
PS	Claim 20; SEQ ID 671; 394pp; English.	
XX		
CC	The present invention relates to novel human coding sequences	
CC	(ABQ9368-ABQ99608) and proteins (ABP6482-ABP65022). The sequences are	
CC	useful in therapeutic, diagnostic and research methods. The	
CC	polynucleotides may be used in the field of molecular biology as	
CC	hybridisation probes, primers for PCR, for chromosome and gene mapping,	
CC	for the recombinant production of protein, or in generation of anti-sense	
CC	DNA or RNA. The polynucleotides are useful in diagnostics as expressed	
CC	sequence tags (ESTs) for identifying expressed genes or for physical	
CC	mapping of the human genome. The proteins may be used as molecular weight	
CC	markers, or as nutritional sources or supplements. The proteins may be	
CC	used to maintain and expand cell population in a totipotential or	
CC	pluripotential state useful for re-engineering damaged or diseased	
CC	tissues, transplantation, manufacture of bio-pharmaceuticals or the	
CC	development of bio-sensors. The polynucleotides and proteins are useful	
CC	for preventing, treating or ameliorating disorders involving aberrant	
CC	protein expression or biological activity, e.g. haematopoietic disorders,	
CC	central/peripheral nervous system diseases, mechanical and traumatic	
CC	disorders, non-healing wounds, immune deficiencies and disorders,	
CC	infectious diseases caused by viral, bacterial or fungal infection,	
CC	autoimmune disorders, allergic reactions and conditions, coagulation	
CC	disorders, or cancer. The polynucleotide sequences of the invention were	
CC	assembled from ESTs isolated mainly by sequencing by hybridisation, and	
CC	in some cases, sequences obtained from one or more public databases.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 404 AA;	
Query Match	7.0%; Score 145.5; DB 23; Length 404;	
Best Local Similarity	23.5%; Pred. No. 0.0012;	

Matches	77; Conservative	34; Mismatches	94; Indels	123; Gaps	15;
QY	64 EPIITNDRTS-----QRYDQGNFT--SEMIHNVFSDSGNIR-----CSLQNRLLHGS	112			
DB	162 KPLVNEKGVSKQETRRHPETGLFTLQSELM---VTPARGDPRPTFCSPGLPRHR	218			
QY	113 AYLTVQVMGELRIP---SVNLVAENEP-----CEVTCPLPSHTWLPDIS	154			
DB	219 ALRTAPIQPRVWEPVPLEVQLV---EPGGAAPGGTIVLTCEVPAQS-----PQIH	270			
QY	155 WELGLIVSHSSYFVPEPSDLSQSAVSIILATPQSNGLTTCVATWKSLLKARKSATVNLTVI	214			
DB	271 WMKD-----GVPPLPPSVLILPEIGPDQGTYSVATHSSHGSPQESRAVSII	321			
QY	215 RCPDQTGGGINIPGVLSLPSLPSLPTWCKVGLGLAGTMLLT-----PCTITIRCC	267			
DB	322 E-PGEEG-----PTAGSVGGSLGTLLALGLGLGTAALLIGVI	361			
QY	268 CRRRCGCCGCCRCFCRRKRGFRIOFKKSEKKT--NKETETESGNENSGYNSDEQ	325			
DB	362 LWQRR-----QRRGERKAPENQEEERAEALN-----	389			
QY	326 KYTDTASLPKSCESDDEQRNNSCOPP	353			
DB	390 -----QSEPEAGESSTGGP	404			
RESULT 13					
AAE23219	AAE23219 standard; Protein; 404 AA.				
XX	AAE23219;				
DT	27-AUG-2002 (first entry)				
XX	Human receptor for advanced glycosylation end product (RAGE) protein.				
XX	Human; Receptor for advanced glycosylation end product; RAGE; cardiant;				
KW	tissue growth; neointimal formation; blood vessel; restenosis; diabetes;				
KW	myocardial infarction; angioplasty; peripheral vascular surgery; angina;				
KW	transgenic animal; acute thrombotic stroke; venous thrombosis.				
XX	Homo sapiens.				
XX	WO200230889-A2.				
XX	18-APR-2002.				
XX	12-OCT-2001; 2001WO-US32036.				
XX	13-OCT-2000; 2000US-0687528.				
XX	(UYCO ) UNIV COLUMBIA NEW YORK.				
XX	Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;				
XX	WPI; 2002-426260/45.				
XX	N-PSDB; AAD36952.				
XX	Inhibiting new tissue growth or neointimal formation in blood vessels				
PT	of subject suffering from diabetes, stroke and preventing restenosis,				
PT	comprises administering inhibitor of receptor for advance glycation end				
XX	product -				
XX	Disclosure; Page 16; 43pp; English.				
XX	The invention relates to a method for inhibiting new tissue growth or				
CC	neointimal formation in blood vessels in a subject that has experienced				
CC	blood vessel injury and preventing exaggerated restenosis in a diabetic				
CC	subject. The method comprises administering an inhibitor of receptor for				
CC	advanced glycation/glycosylation end product (RAGE), so as to inhibit new				
CC	tissue growth or neointimal formation in subject's blood vessels and				
CC	preventing restenosis in the subject. The method is useful for inhibiting				

new tissue growth or neointimal formation in blood vessels in a subject like non-human animal, a transgenic non-human animal or a human suffering from diabetes, acute thrombotic stroke, venous thrombosis, unstable angina, myocardial infarction, abrupt closure following angioplasty or stent placement, or thrombosis as a result of peripheral vascular surgery. The method is also useful for preventing restenosis and for determining whether a compound inhibits new tissue growth in a blood vessel in a subject. The present sequence is human receptor for advanced glycosylation end product (RAGE) protein.

XX Ameliorating neurovascular stress and decreasing cerebral  
PT vasoconstriction in subject suffering from chronic/acute cerebral  
PT amyloid angiopathy, by administering inhibitor of receptor for advanced  
PT glycation endproduct -  
PT  
PS Disclosure; Page 16; 68pp; English.  
XX  
XX The invention describes a method of ameliorating neurovascular stress,  
CC and decreasing cerebral vasoconstriction in subject suffering from  
CC chronic or acute cerebral amyloid angiopathy, comprising administering  
CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).  
CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain  
CC barrier, thus decreasing cerebral vasoconstriction and increasing  
CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a  
CC subject, decreasing cerebral vasoconstriction in a transgenic non-human  
CC animal (preferably, transgenic mouse overexpressing mutant human amyloid  
CC beta precursor protein) or a human, suffering from chronic or acute  
CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for  
CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy  
CC in a subject, where the neurovascular stress is caused by Alzheimer's  
CC disease, aging, Down's syndrome, head trauma or stroke. This is the  
CC amino acid sequence of human receptor for advanced glycation end  
CC product (RAGE) described in the invention.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:28:26 ; Search time 15.3865 Seconds  
(without alignments)  
2412.580 Million cell updates/sec

Title: US-09-729-264-4

Perfect score: 2088

Sequence: 1 MVAGAMENRDPGSGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	145.5	7.0	404	1 I61596	advanced glycosylation
2	136	6.5	1447	2 A54100	tumor suppressor p
3	134.5	6.4	1427	2 I51669	tumor suppressor -
4	133.5	6.4	871	1 I48696	protein-tyrosine k
5	133.5	6.4	881	1 I48697	protein-tyrosine k
6	129.5	6.2	164	2 T24272	hypothetical prote
7	122.5	5.9	188	2 T15651	hypothetical prote
8	121	5.8	344	2 I56551	neurotrophin - rat
9	119.5	5.7	6642	2 T29757	protein UNC-89 - C
10	119	5.7	1070	2 J04593	protein-tyrosine k
11	115.5	5.5	946	1 A47299	ror-related recept
12	114.5	5.5	152	2 T18975	hypothetical prote
13	114.5	5.5	531	2 S20900	titin - mouse (fra
14	114.5	5.5	862	2 I49583	differentiation an
15	114.5	5.5	28926	1 I38344	titin, cardiac mus
16	112.5	5.4	345	2 J04025	opioid-binding cel
17	112.5	5.4	416	1 A42879	advanced glycosyla
18	112	5.4	868	2 A46512	CD22 homolog/B lym
19	111.5	5.3	345	2 S03199	opioid-binding pro
20	111.5	5.3	1443	2 I50600	neogenin - chicken
21	111.5	5.3	6805	2 S20901	titin - rabbit (fr
22	111	5.3	620	2 JH0593	Schwann cell myeli
23	111	5.3	693	2 S49228	sodium-dependent p
24	110.5	5.3	338	2 J01238	opioid-binding pro
25	110.5	5.3	345	2 J01239	opioid-binding pro
26	110.5	5.3	4391	2 A38096	perlecan precursor
27	109.5	5.2	364	2 A30521	myeloid cell surfa
28	109	5.2	3375	2 T19821	hypothetical prote
29	107	5.1	391	2 T09058	butyrophilin homol

30	106.5	5.1	487	2 S65133	butyrophilin - mou
31	106.5	5.1	802	2 T13149	mitogen-and stress
32	106.5	5.1	841	2 JC5894	killer cell inhibi
33	106.5	5.1	1272	2 S26180	neurofascin - chic
34	106	5.1	1092	1 JN0635	neural cell adhesi
35	105.5	5.1	662	2 T16525	hypothetical prote
36	104.5	5.0	1177	2 T16594	hypothetical prote
37	104	5.0	764	2 A49448	irregular chiasm C
38	104	5.0	2325	2 C88369	protein unc-52 [im
39	104	5.0	5825	2 T12117	polyprotein - fava
40	103.5	5.0	2491	1 A28372	insulin-like growt
41	103.5	5.0	3707	2 S18252	heparan sulfate pr
42	102.5	4.9	890	1 A53743	protein-tyrosine k
43	102	4.9	721	2 T41530	hypothetical prote
44	101.5	4.9	215	2 T16542	hypothetical prote
45	101.5	4.9	1449	2 S47423	E2 glycoprotein pr

## ALIGNMENTS

### RESULT 1

I61596

Advanced glycosylation end-products receptor precursor - human  
N/Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprote  
C/Species: Homo sapiens (man)  
C/Date: 24-May-1996 #sequence revision 07-Feb-1997 #text\_change 16-Jul-1999  
C/Accession: I61596; B42879; S27968  
R/Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko, H  
Genomics 23, 408-419, 1994  
A/Title: Three genes in the human MHC class III region near the junction with the clas  
nterpart of mouse mammary tumor gene int-3.  
A/Reference number: A55562; MUID:95137587; PMID:7835890  
A/Accession: I61596

A/Status: nucleic acid sequence not shown; translation not shown; translated from GB/EI  
A/Molecule type: DNA  
A/Residues: 1-404 <RES>  
A/Cross-references: GB:D28769; NID:9561657; PIDN:BAA05958.1; PID:9561659  
R/Neepar, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; J  
J. Biol. Chem. 267, 14998-15004, 1992  
A/Title: Cloning and expression of a cell surface receptor for advanced glycosylation  
A/Reference number: A42879; MUID:92340547; PMID:1378843  
A/Accession: B42879

A/Molecule type: mRNA  
A/Residues: 'G', '2-99', 'R', '101-404 <NEE>  
A/Cross-references: EMBL:M91211; NID:9190845; PIDN:AAA03574.1; PID:9190846  
A/Experimental source: lung  
A/Note: sequence extracted from NCBI backbone (NCBIP:109438)  
C/Comment: Advanced glycosylation end products are heterogeneous nonenzymatically glyco  
cellular function, thus contributing to tissue lesions in diabetes.  
C/Comment: This receptor appears also to mediate the effects of amyloid beta peptide on  
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.  
C/Genetics:  
A/Gene: GDB:AGER  
A/Cross-references: GDB:306354; OMIM:600214  
A/Map position: 6p21.3-6p21.3  
A/Introns: 18/1; 53/3; 119/1; 140/3; 170/1; 231/1; 274/3; 322/1; 331/1; 373/2  
C/Function:  
A/Description: neuronal receptor for amphoterin, a DNA-binding protein involved in neu:  
C/Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
C/Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein  
F/1-22/Domain: signal sequence #status predicted <SIG>  
F/23-404/Domain: advanced glycosylation end products receptor #status predicted <NAT>  
F/23-344/Domain: extracellular #status predicted <EXT>  
F/31-101/Domain: immunoglobulin homology <IM1>  
F/137-210/Domain: immunoglobulin homology <IM2>  
F/252-303/Domain: immunoglobulin homology <IM3>  
F/345-362/Domain: transmembrane #status predicted <TMW>  
F/363-404/Domain: intracellular #status predicted <INT>  
F/25\_81/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/38-99,144-208,259-301/Disulfide bonds: #status predicted

Query Match

7.0%; Score 145.5; DB 1; Length 404;

RESULT 3  
I51669  
tumor suppressor - African clawed frog  
C\_Species: Xenopus laevis (African clawed frog)  
C\_Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C\_Accession: I51669  
R\_Pieceall: W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon,  
Dev. Biol. 166, 654-665, 1994  
A\_Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene  
A\_Reference number: I51668; MUID:95113183; PMID:7813784  
A\_Accession: I51669  
A\_Status: preliminary; translated from GB/EMBL/DBJ  
A\_Molecule type: mRNA  
A\_Residues: 1-427 <PIE>  
A\_Cross-references: EMBL:U0986; NID:9606873; PID:AA70168.1; PID:9606874  
C\_Genetics:  
A\_Gene: XDCCa

Query Match	6.4%;	Score 134.5;	DB 2;	Length 1427;
Best Local Similarity	25.4%;	Pred. No. 0.065;		
Matches	65;	Conservative 33;	Mismatches 93;	Indels 65; Gaps 15
QY	12	PGSGS-GNEV-----	IEGQNARVLKGSQARENCVTSQGW-KLIMNAL	52
DB	217	PGSARVGNEAEILRILSESGLHRQCVFLQPSNVVAIEGGQDAVLECAVS-GYPTPTIVMVG	275	
QY	53	SDMVVLSVRPEPIITNDRFTSQRYDOGGNFTSEMIHNHVEPDSGNIRSC-SLONSRLH	110	
DB	276	GD-----EPVPIRTR---KYSVLGG--SNLLISNVDDDDAGAYTCVATYKNENTS	320	
QY	111	GSAYILTVOVMGELFIPSNVLVAENEPCEVTCL---PSHWTWLPDISW-ELGLLVSHSS	165	
DB	321	PSADULTVMVPQFLNHPANLYAYESMDIEPECAVSGKPS-----PTVKWTKNGEVIIPSD	375	
QY	166	YFVFEPSDILQASVILALTPQSGNLTLCVA-----TWKSIAKARKSATNLNLTVI-RC	216	
DB	376	YFQIVDGSNLR-----ILGLVKSDEGYQCIAREAGNIQYAOILTPDPAPVSSSILPSA	431	
QY	217	PDQTGGGINIPGVLS	232	
DB	432	PRDV-----VPVLVSS	442	

RESULT 4  
I48696  
protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 2 - mouse  
N;Alternate names: receptor-type tyrosine kinase  
N;Contains: protein-tyrosine kinase nsk2 precursor, splice form 4  
C;Species: Mus musculus (house mouse)  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000  
C;Accession: I48696; S60738  
R;Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D.  
Oncogene 11, 281-290, 1995  
A;Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosi  
A;Reference number: I48696; MUID:95349951; PMID:7624144  
A;Accession: I48696  
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-871 <NAME>  
A;Cross-references: EMBL:X86444; NID:g929723; PIDN:CAA60165.1; PID:g929724  
A;Experimental source: splice form 2  
A;Accession: S60738

C:Comment: For alternate splice forms see PIR:I48696.

C:Genetics:

A:Gene: nsk2

A:Cross-references: MGI:103308

C:Superfamily: mouse ror-related receptor; immunoglobulin homology; protein kinase homology

C:Keywords: ATP; Glycophenate; phosphotransferase; receptor; transmembrane protein; tyrosine kinase

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-881/Product: protein-tyrosine kinase nsk2, splice form 1 #status predicted <MARI>

F:22-456,'A',466-881/Product: protein-tyrosine kinase nsk2, splice form 3 #status predicted <MARI>

F:42-101/Domain: immunoglobulin homology <IMM1>

F:135-192/Domain: immunoglobulin homology <IMM2>

F:226-284/Domain: immunoglobulin homology <IMM3>

F:498-518/Domain: transmembrane #status predicted <TRM>

F:575-865/Domain: protein kinase homology <KIN>

F:583-591/Region: protein kinase ATP-binding motif

F:282,462/Binding site: carboxyphate (Asn) (covalent) #status predicted

Query Match 6.4%; Score 133.5; DB 1; Length 881;

Best Local Similarity 21.6%; Pred. NO. 0.046;

Matches 73; Conservative 45; Mismatches 127; Indels 93; Gaps 17;

[illegible]

```

C;Species: caenorhabditis elegans
C;date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24272
R;Sims, M.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z19867
A;Accession: T24272
A;status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-164 <will>
A;Cross-references: EMBL:Z66499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:T01B7
A;Experimental source: clone T01B7
C;Genetics:
A;Gene: CESP:T01B7.8
A;Map position: 2
A;Introns: 20/3; 90/2

Query Match 6.2%; Score 129.5; DB 2; Length 164;
Best Local Similarity 32.1%; Pred.No 0.015;
Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps

QY 157 LGLLVSHSHSYFVPEPSPDQASVILALTQSQNGTLTCVATWKLKARKSATVNLTVIRC :
Db 6 LAIIAIGTFTAV----SQVQSAV-----LPVSTFELATVCTDYSTASTADTGTGSSSSSV :

```



Db 21 LQLLTSGNGKSDFLVGLPPLAIVGQDKELPKCLSLNLSAEGMELRWYRDXSPVVHV 80  
 Qy 57 RMEPIIINDRTSRYOQGNFTSEMI-----IHVPSDSGNIRCLQNSRLHGS 108  
 Db 81 YKNGEDVYDEQMV--EYKRTSFGNSHARGAAVKIHNVTVDNGTVCHVKFYKTSYHSQ 138  
 Qy 109 AVLTQVMGELFIPSNLVVAENECVETCLPSHWRTPDISW-EL-GLLVSHSSYVFPV 166  
 Db 139 ATLWIKVAGSGSPRIRVTDQKIRAECTSGAGWPKYKVENLDLKGQVSAESHFSVS 198  
 Qy 167 EPSDIQSASVILALPQSN--GTLTCVATWKLKARKSATVNL-TVIRCPQDTGGGINIP 223  
 Db 199 ASTGLVALLSI--VTPQDTAVGGLTCSISNPLLPKVTETVYLLASLSRPLSTESGPAIP 256  
 Qy 224 GVLSSIPSLGFLPTMGKVLGLAGLMLTPTCTLTIRCCGRRCCGCGCCCFPCR 283  
 Db 257 LILTAL-----GLVSAIAA-----CAFGK----- 276  
 Qy 284 RKRFRIOQKSEKTKETTESGNSG--YNSDEQKTTDTASLPKSCSSDPDQEQ- 341  
 Db 277 -----HKEHKTQNEEDPGARDEAGLHVLSLDPETAS--PKLMVSDQKSV 323  
 Qy 342 ----RNSCCGPHQADQRP 357  
 Db 324 KRLFPDQVFPSSRRFNQDP 343

RESULT 4  
 Q9D912 PRELIMINARY; PRT; 168 AA.  
 ID Q9D912 (TEMBLrel. 17, Created)  
 AC Q9D912 (TEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TEMBLrel. 22, Last annotation update)  
 DE 4931420D14RIK protein.  
 GN 4931420D14RIK  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; Tissue=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Hoffell D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,  
 RA Hayaishizaki Y.,  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL MEDL: AK006892; BAB24782.1; -;  
 DR MGD; MGI:1913992; 4931420D14RIK  
 SQ SEQUENCE 168 AA; 18931 MW; 7A2BD279612ASE94 CRC64;

Query Match 6.8%; Score 141; DB 11; Length 168;  
 Best Local Similarity 32.0%; Pred. No. 3.3e-05;  
 Matches 39; Conservative 20; Mismatches 49; Indels 14; Gaps 6;  
 Qy 256 CTITIRCCGRRCCGNC--CCRCFCRRKGRF-----IQFKKSEKTKETTESG 310

Db 55 CSLSRSCCCRCRC--CHCRCRC--CCSRRRFRSRTTLKFFQITEKGEQSLQRRIRQ 111  
 Qy 311 NENSGYNSDEQKTTDTASLPKSCSSDPEQRNSCGPHQADQRPSPASHQASPNL 370  
 Db 112 LTRQLELIEPE--PTMALEPSBITVAFFSHKNANVSDP-----EVPFCLDSDFPFGDL 165  
 Qy 371 AS 372  
 Db 166 AS 167

RESULT 5  
 Q8CH20 PRELIMINARY; PRT; 168 AA.  
 ID Q8CH20 (TEMBLrel. 23, Created)  
 AC Q8CH20 (TEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)  
 DE Basic protein CKTIR3.  
 GN CKTIR3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; Tissue=Testis;  
 RA Xu X., Bai X., Silvius D., Escalier D., McFarland L., Xu P.-X.;  
 RT "CK2 differentially phosphorylate a family of novel spermatid-specific  
 RT basic nuclear proteins."  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF463502; AAO15675.1; -;  
 SQ SEQUENCE 168 AA; 18957 MW; 8F30D3D27B9BF595 CRC64;

Query Match 6.7%; Score 138; DB 11; Length 168;  
 Best Local Similarity 32.0%; Pred. No. 6.4e-05;  
 Matches 39; Conservative 19; Mismatches 50; Indels 14; Gaps 6;  
 Qy 256 CTITIRCCGRRCCGNC--CCRCFCRRKGRF-----IQFKKSEKTKETTESG 310  
 Db 55 CSLSRSCCCRCRC--CHCRCRC--CCSRRRFRSRTTLKFFQITEKGEQSLQRRIRQ 111  
 Qy 311 NENSGYNSDEQKTTDTASLPKSCSSDPEQRNSCGPHQADQRPSPASHQASPNL 370  
 Db 112 LTRQLELIEPE--PTMALEPSBITVAFFSHKNANVSDP-----EVPFCLDSDFPFGDL 165  
 Qy 371 AS 372  
 Db 166 AS 167

RESULT 6  
 Q9D4K2 PRELIMINARY; PRT; 173 AA.  
 ID Q9D4K2 (TEMBLrel. 17, Created)  
 AC Q9D4K2 (TEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TEMBLrel. 22, Last annotation update)  
 DE 4931420D14RIK protein.  
 GN 4931420D14RIK  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; Tissue=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Hoffell D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,  
 RA Hayaishizaki Y.,  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL MEDL: AK006892; BAB24782.1; -;  
 DR MGD; MGI:1913992; 4931420D14RIK  
 SQ SEQUENCE 168 AA; 18931 MW; 7A2BD279612ASE94 CRC64;



```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=95113183; PubMed=7813784;
RA  Pierceall W.E., Reale M.A., Candia A.P., Wright C.V., Cho K.R.,
RT  Fearon E.R.;
RT  "Expression of a homologue of the deleted in colorectal cancer (DCC)
RT  gene in the nervous system of developing Xenopus embryos.";
RL  Dev. Biol. 166:654-665(1994).
DR  EMBL; U10986; AAA70168.1; -.
DR  HSSP; P40189; 1BQV.
DR  InterPro; IPR003962; FNIII subd.
DR  InterPro; IPR003961; FN III.
DR  InterPro; IPR007110; Ig-like.
DR  InterPro; IPR003598; Ig_c2.
DR  InterPro; IPR003006; Ig_MHC.
DR  Pfam; PF00041; fn3; 6.
DR  Pfam; PF00047; Ig; 4.
DR  PRINTS; PR00014; FNTYPEIII.
DR  SMART; SM00060; FN3; 6.
DR  SMART; SM00408; IGC2; 3.
DR  SMART; PS50835; IG LIKE; 4.
KW  Immunoglobulin domain; Repeat.
SQ  SEQUENCE 1427 AA, 156533 MW, 61FEA12C8A674972 CRC64;

Query Match      6.5%; Score 133; DB 13; Length 1427;
Best Local Similarity 25.7%; Pred. No. 0.0032;
Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY  17 IEQPQNAVILKGSQARFNCTVSGW--KLIMWALSDMVLSVRPMEPIITNDRFTSQRYD 74
    :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  243 LORPSNVVIEGQDAVLECAVS-GYETPTIVMVGQD-----EPVPIRTR---KYS 288
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  75 QGNGFTSEMIHNVPSDSGNIRC--SLQNSRLHGSAYLTQVMGELFIPSNLVVAENE 132
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  289 VLGG--SNLLISNVTDGAGYTCVATYKNTSFSADLTVMVPPQFLNHPANLYAYESM 346
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  133 PCEVTCL----PSHWTRLPDISW-ELGLVSHSSYFVPEPDLQSAVILALTQSNGT 187
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  347 DIEFECAVSGKPS-----PVRKTKNGEVVPSDFQIVDGSNLR----ILGLVKSDEGY 397
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  188 LTCVA-----TWKSLKARKSATVNLTVI-RCPODTGGGGINIPGVLS 228
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  398 YQCIANEAGNIQTYAQLIIPDPAPVSSSILPSAPRDV-----VPVLVSS 442
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 9
Q9DF61 ID Q9DF61 PRELIMINARY; PRT; 344 AA.
AC Q9DF61;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OSCAM alpha 1 isoform.
GN OSCAM.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN SEQUENCE FROM N.A.
RP MEDLINE=20499204; PubMed=11042360;
RA Lodge A.P., Howard M.R., McNamee C.J., Moss D.J.;
RT "Co-localisation, heterophilic interactions and regulated expression
RT of IgGon family proteins in the chick nervous system.";
RL Brain Res. Mol. Brain Res. 82:84-94(2000).
DR EMBL; AF292334; AAG01877.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; IGC2; 2.

DR PROSITE; PS50835; IG LIKE; 3.
KW Immunoglobulin domain.
SQ SEQUENCE 344 AA, 37531 MW, 37FE6051CBF0E7B4 CRC64;

Query Match      6.4%; Score 132.5; DB 13; Length 344;
Best Local Similarity 25.5%; Pred. No. 0.00054;
Matches 70; Conservative 38; Mismatches 104; Indels 63; Gaps 16;

QY  4 VIFLHSGS-----SGNEVI-EGPQNAVILKGSQARFNCTVSGWKLIMWALSDMVLS 54
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  21 LLFLVPAGVPVRSRGTATFPKAMDNVTVRQGESATLRCCTVDORVRRVAV-LNRSTILYAGN 79
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  55 ---SVRPMPEIITNDRFTSQRYDQGNGFTSEMIHNVPSDSGNIRCQLQ-----NSRL 105
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  80 DKWSIDNRVILSN--TKQY-----SIKHNVVDVDEGPTCSVQTDNHPKTSRV 128
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  106 HGSAYLTQVMGELFIPSNLVVAENPCEVTCL----PSHWTRLPDISWELGLLVSHSS 161
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  129 H---LIVQVPPQIVNISSDITVNEGSSVTLMLCLAFGRPE-----PTVTR--HLSGKG 176
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  162 YFVPEPDLQSAVILALTQSNGTLTCTVATWKSARKSATVNLTV-----IRCPQDT 216
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  177 QGFVSEDEYLE-----ITGITREQSGEYEC SAV-NDNAVDPVRKVKTVNYPPIISNAKNT 231
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  217 GGGINIPGVL-----SSLPSLGFSLPTWKGVLGLA 247
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  232 GASVQKGILQCEASAVPVAEPQ---WFKEDTRLA 263
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 10
Q9Z290 ID Q9Z290 PRELIMINARY; PRT; 697 AA.
AC Q9Z290;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type IIB Na/phosphate-cotransporter.
GN SLC34A2 OR NPT2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
SEQUENCE FROM N.A.
RP STRAIN=NMRI; TISSUE=Small intestine;
RX MEDLINE=99045724; PubMed=9826740;
RA Hilfinger H., Hattenhauer O., Traebert M., Forster I., Murer H.,
RA Biber J.;
RT "Characterization of a murine type II sodium-phosphate cotransporter
RT expressed in mammalian small intestine.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14564-14569(1998).
DR EMBL; AF081499; AAC80007.1; -.
DR MGD; MGI:1342284; Slc34a2.
DR InterPro; IPR003841; Na/Pi cotranspt.
DR Pfam; PF02690; Na Pi cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
SQ SEQUENCE 697 AA, 76286 MW, 839B5CCB0F565265 CRC64;

Query Match      6.4%; Score 131; DB 11; Length 697;
Best Local Similarity 20.6%; Pred. No. 0.0019;
Matches 89; Conservative 51; Mismatches 123; Indels 168; Gaps 23;

QY  3 LVIFLHSGSGNEVIEGPQNAVILK-----GSOARENCTVSQ 40
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  382 MIVKLLGS-----VLRG-QVATVIKKTINTDFFPPFAWLTYLALYGAGMTFIVQSSV 435
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  41 WKLIWALSDMVLSVRPMEPIITNDRFTSQRYDQGNGFTSEMIHNVPSDSGNIRC 100
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  436 FTSAMTPLIGIVISIERAYPLTGSNI-----GTTTALLAALASFGNT--LASSL 485
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY  101 QNSRLHGSAYLTQVMGELFIPSNLVVAENPCEVTCLPSHWTRLP-----DIS-- 150
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  486 QIALCH-----FFFNISGI-----LLWYPIPTPLRLAKGLGNISAK 524
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```



```

Db      58 KRQGGCGCGCGC-----GCCGGGGGGG--CGCCCRPRCCCCCRCTC 101
QY      273 --NCCC-RCFCOR 283
Db      102 CRTCCCTRCCTC 115

RESULT 13
Q9DGI5 ID Q9DGI5 PRELIMINARY; PRT; 315 AA.
AC Q9DGI5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CEPU-Se alpha 1 isoform.
GN CEPU-SE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lodge A.P., McNamee C.J., Howard M.R., Reed J.E., Moss D.J.;
RT "Characterisation of CEPU-Se, a secreted isoform of the IGLON family
RL protein CEPU-1";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292336; AAG01879.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 3.
DR SMART; SM00408; IGC2; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
KW Immunoglobulin domain
SQ SEQUENCE 315 AA; 34606 MW; 58C5D27F0DDC6F62 CRC64;

Query Match 6.2%; Score 127.5; DB 13; Length 315;
Best Local Similarity 25.4%; Pred. No. 0.0014;
Matches 68; Conservative 41; Mismatches 98; Indels 61; Gaps 15;

QY 4 VIFLHGSG-----SGNEVI-EGPONATVLKGSQARFNCTVSQ-----GW---KLIMWALSD 50
Db 21 LLFLVAGVPVRSGDATFPKAMDNVTVROGESATLRCSDVNRVTRVAVLNRSSILYAGND 80
QY 51 MVVLSVRPMEPIITNDRFTSQDYDQGNFTSEMIHNVPSPDSGNTRCSLQ-----NSR 104
Db 81 KWCLDPRVLLANTKTOYSIQ-----IHDVDYDEGPTYCSVQTDNHPKTSR 127
QY 105 LHGSAYLTQVMGELFTPSVNLVAENEPCEVTCPLFSHWTRLPD--ISWELGLLYSHSY 162
Db 128 VH-----LIVQSPKITEISSDISINEGNNVSLTCLIA---TGRDPTITWR---HISPRKV 177
QY 163 YFVPEPSDLSQSAVSIALTQPSNGTLTCVATWKLKARKSATVNLTV-----IRCPQDTG 217
Db 178 GFISEDEYLE-----ITGITREQSGEYECSSAS-NDVAAPVQVRVKVTVNPPYISDAKSTG 232
QY 218 GGINIPGVL-----SSLPSLGLFSLPTWCK 241
Db 233 VPVGQKGLMCEASAVPSADFQ---WYK 257

RESULT 14
Q93242 ID Q93242 PRELIMINARY; PRT; 344 AA.
AC Q93242;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CEPU-1.
OS Gallus gallus (Chicken).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Kimura Y., Shirabe K., Fukushima M., Takeshita M., Tanaka H.;
RT "CEPU-1: an Immunoglobulin Superfamily Molecule, Has Cell Adhesion
RT Activity and Shows Dynamic Expression Patterns in Chick Embryonic
RT Spinal Cord.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011810; BAA31514.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 3.
DR SMART; SM00408; IGC2; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
KW Immunoglobulin domain.
SQ SEQUENCE 344 AA; 37613 MW; 22CAA8F526A6B57E CRC64;

Query Match 6.2%; Score 127.5; DB 13; Length 344;
Best Local Similarity 25.4%; Pred. No. 0.0016;
Matches 68; Conservative 41; Mismatches 98; Indels 61; Gaps 15;

QY 4 VIFLHGSG-----SGNEVI-EGPONATVLKGSQARFNCTVSQ-----GW---KLIMWALSD 50
Db 21 LLFLVAGVPVRSGDATFPKAMDNVTVROGESATLRCSDVNRVTRVAVLNRSSILYAGND 80
QY 51 MVVLSVRPMEPIITNDRFTSQDYDQGNFTSEMIHNVPSPDSGNTRCSLQ-----NSR 104
Db 81 KWCLDPRVLLANTKTOYSIQ-----IHDVDYDEGPTYCSVQTDNHPKTSR 127
QY 105 LHGSAYLTQVMGELFTPSVNLVAENEPCEVTCPLFSHWTRLPD--ISWELGLLYSHSY 162
Db 128 VH-----LIVQSPKITEISSDISINEGNNVSLTCLIA---TGRDPTITWR---HISPRKV 177
QY 163 YFVPEPSDLSQSAVSIALTQPSNGTLTCVATWKLKARKSATVNLTV-----IRCPQDTG 217
Db 178 GFISEDEYLE-----ITGITREQSGEYECSSAS-NDVAAPVQVRVKVTVNPPYISDAKSTG 232
QY 218 GGINIPGVL-----SSLPSLGLFSLPTWCK 241
Db 233 VPVGQKGLMCEASAVPSADFQ---WYK 257

RESULT 15
Q61987 ID Q61987 PRELIMINARY; PRT; 871 AA.
AC Q61987;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nsk2 protein precursor.
GN MUSK OR Nsk2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP TISSUE=Myoblasts;
RX MEDLINE=95349951; PubMed=7624144;
RA Ganju P., Walls E., Brennan J., Reith A.D.;
RT "Cloning and developmental expression of Nsk2, a novel receptor
RT tyrosine kinase implicated in skeletal myogenesis,.";
RL Oncogene 11:281-290(1995).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X86444; CAA60165.1; -.
DR HSSP; P11362; IFCK.
DR M3D; MGI:103581; Musk.
DR InterPro; IPR002453; Beta tubulin.
DR InterPro; IPR000024; Pz_domain.

```

Search completed: November 13, 2003, 03:33:34  
Job time : 38.0976 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:31:20 ; Search time 12.2478 Seconds  
(without alignments)  
1319.641 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059

Sequence: 1 MGLVFLHSGSGGNEVIEGP.....HPQASFNLAGPEKVSNTTVV 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pap.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pap.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pap.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pap.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pap.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	147.5	7.2	404	US-09-638-649-3	Sequence 3, Appli
2	128.5	6.2	1345	US-08-977-767-3	Sequence 3, Appli
3	128	6.2	1447	US-09-041-886-25	Sequence 25, Appli
4	128	6.2	1447	PCT-US94-05277-2	Sequence 2, Appli
5	126.5	6.1	869	US-08-374-834-16	Sequence 16, Appli
6	126.5	6.1	869	US-08-644-271-29	Sequence 29, Appli
7	126.5	6.1	869	US-09-077-955-33	Sequence 33, Appli
8	123.5	6.0	332	US-08-062-365-1	Sequence 1, Appli
9	122.5	5.9	340	US-09-651-200-2	Sequence 2, Appli
10	122.5	5.9	441	US-09-651-200-4	Sequence 4, Appli
11	121.5	5.9	534	US-09-651-200-6	Sequence 6, Appli
12	121.5	5.9	534	US-09-651-200-24	Sequence 24, Appli
13	118.5	5.8	318	US-08-633-148-4	Sequence 4, Appli
14	118.5	5.8	340	US-08-633-148-2	Sequence 2, Appli
15	117	5.7	325	US-09-651-200-20	Sequence 20, Appli
16	115	5.6	352	US-09-996-243-505	Sequence 505, App
17	114	5.5	1395	US-09-540-245A-15	Sequence 15, Appli
18	113.5	5.5	416	US-09-638-649-1	Sequence 1, Appli
19	112	5.4	868	US-08-374-834-1	Sequence 1, Appli
20	112	5.4	868	US-08-644-271-1	Sequence 1, Appli
21	112	5.4	868	US-09-077-955-1	Sequence 1, Appli
22	110	5.3	689	US-09-499-964-1	Sequence 1, Appli
23	109.5	5.3	478	PCT-US95-08493-15	Sequence 15, Appli
24	109.5	5.3	860	PCT-US95-08493-19	Sequence 19, Appli
25	109.5	5.3	868	PCT-US95-08493-21	Sequence 21, Appli
26	107.5	5.2	362	US-08-415-751-6	Sequence 6, Appli
27	107	5.2	365	US-08-979-424-3	Sequence 3, Appli

## RESULT 1

US-09-638-649-3

; Sequence 3, Application US/09638649

; Patent No. 6563015

; GENERAL INFORMATION:

; APPLICANT: Stern, David M.

; APPLICANT: Schmidt, Ann Marie

; APPLICANT: Fan, Shi Du

; TITLE OF INVENTION: TRANSGENIC MICE OVER-EXPRESSING RECEPTOR FOR ADVANCED

; TITLE OF INVENTION: GLYCATION ENDPRODUCT (RAGE) AND MUTANT APP IN BRAIN AND

; FILE REFERENCE: 0575/62175

; CURRENT APPLICATION NUMBER: US/09/638,649

; CURRENT FILING DATE: 2000-08-14

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 3

; LENGTH: 404

; TYPE: PRT

; ORGANISM: Human

US-09-638-649-3

Query Match 7.2%; Score 147.5; DB 4; Length 404;

Best Local Similarity 22.6%; Pred. No. 4.6e-05;

Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

Qy 15 EVIEGPQATVLKGSQARENCTVSQ---WKLMWLSDMVLSVRPMEPIITNDRETS- 70

Db 125 EIVDSASELTA--GVPNKVGTCVSGSPAGTSLSHLDG-----KPLVNEKGVSV 173

Qy 71 ----QRYDGGNFT--SEMIHNVEPDSGNTR---CSLQNSRLHGSAYLTVQVMGELF 120

Db 174 KEQTRRHPTGLTLOSELM--VTPARGDPRPTFSCFSFGLPRHRLARTAPIQPRVW 230

Qy 121 IP-----SVNLVVAENP-----CEVCLSHWTRLPDISWELGLLVSHSY 162

Db 231 EPVPLEVOLVV---EPEGGAVAPGTVTLTCEVPAQPS-----FQIHWMDK----- 274

Qy 163 YFVPEPDDIQQSAVSIHALTPQSGNTLTCTVATWKSLSKARSAFVNLTVTRCPDGTGGINI 222

Db 275 -GVPLPSPVLIIPETGPOQGYSCVATHSSHPQESRAVSISITB-PEEG----- 327

Qy 223 PGVLSLPSLPSLPTWGVGLAGTMLLT-----PTCTLTIRCCCRRCGCCNCC 275

Db 328 -----PTAGSVGSGGLTALALGILGGLGTALLICVILWQRR----- 366

Qy 276 CRCCFCRRKRGRIQFOKKSEKXT--NKEPETESNGNSGYNSEDEQKTTTASLPKPS 333

Db 367 -----QRRGERKAPNQEEERAEIN----- 389

## ALIGNMENTS

28	107	5.2	365	3	US-09-272-496-2	Sequence 2, Appli
29	106.5	5.2	946	5	PCT-US95-08493-13	Sequence 13, Appli
30	106	5.1	319	1	US-08-997-495B-22	Sequence 22, Appli
31	106	5.1	319	3	US-09-068-051A-22	Sequence 22, Appli
32	106	5.1	319	4	US-09-336-536-67	Sequence 67, Appli
33	106	5.1	319	4	US-09-254-465A-6	Sequence 6, Appli
34	106	5.1	801	1	US-07-906-349A-6	Sequence 6, Appli
35	105	5.1	285	3	US-08-482-085B-20	Sequence 20, Appli
36	105	5.1	1297	3	US-09-540-245A-17	Sequence 17, Appli
37	105	5.1	1381	3	US-09-540-245A-16	Sequence 16, Appli
38	104.5	5.1	879	1	US-08-554-612C-1	Sequence 1, Appli
39	104	5.1	421	2	US-08-659-984A-1	Sequence 1, Appli
40	104	5.1	421	3	US-08-660-531-1	Sequence 1, Appli
41	104	5.1	444	2	US-08-659-984A-5	Sequence 5, Appli
42	104	5.1	444	3	US-08-660-531-5	Sequence 5, Appli
43	103.5	5.0	890	1	US-08-445-640-2	Sequence 2, Appli
44	103.5	5.0	890	3	US-08-170-558-2	Sequence 2, Appli
45	103.5	5.0	890	3	US-08-447-314-2	Sequence 2, Appli

Qy 334 CESSDPEORNSCGPP 349  
Db 390 -QSEEPAGESSTGGP 404

## RESULT 2

US-08-977-767-3  
; Sequence 3, Application US/08977767  
; Patent No. 5972884  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Greenwald, Sara  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: CARBONIC ANHYDRASE VIII  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,767  
; FILING DATE: Herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0423 US  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1345 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1532042

Query Match 6.2%; Score 128.5; DB 2; Length 1345;  
Best Local Similarity 33.0%; Pred. No. 0.012;  
Matches 37; Conservative 1; Mismatches 41; Indels 33; Gaps 5;

Qy 186 GTLTCVATWSLKARKSATVNLTVIRCPQDTGGI-----NIPGVLSLPSLGFSLPTWGX 241  
Db 414 GTCTCTGT-----GC-CGTGGGAAGCGTCAGAGCCCGCTGGATGTGA 455  
Qy 242 VGLGLAGTMTLT-PTCLTIRCCCRRCRCNCCRC-----CFCC 282  
Db 456 CGTGAAGAGGTCTCTATGACCCCTCTCTGCCCCCTCTGAGACTCAGACC 507

## RESULT 3

US-09-041-886-25  
; Sequence 25, Application US/09041886  
; Patent No. 6235872  
; GENERAL INFORMATION:  
; APPLICANT: Bredesen, Dale E.

; APPLICANT: Rabizadeh, Sharroz  
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence  
; TITLE OF INVENTION: Polypeptides and Methods of Use  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/041.886  
; FILING DATE:

; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2626  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-041-886-25

Query Match 6.2%; Score 128; DB 3; Length 1447;  
Best Local Similarity 24.8%; Pred. No. 0.015;  
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

Qy 10 SSGSGNEV-----IEGPNATVVLKGSQARFNCTVSQGW--KLIMWALS DM 51  
Db 220 SRTGNEAEVRILSDPLHRLQLYFLQPSNVVAIEGKDAVLECCVS-GYPPPSFTWLRGE 278  
Qy 52 VVLSVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVEPSDSGNIRC--SLQNSRLHSGA 109  
Db 279 VI-----QLRSKYSLLGG--SNLLISNVTTDDSGMYTCVVTYKNENISASA 323  
Qy 110 YLTVOVMGSELFIPSVNLVVAENEPCEVTCLPSHWTRLPDISM-ELGLLVSHSSYVFVPEP 168  
Db 324 ELTVLVPFNLNHPNLSLYAYESMDIEFECTVS-GKPVTVNMKNGDVVIFSDYFQIVGG 382  
Qy 169 SLDQSAVSILALTPQSNGLTLCVATWKSILKARKSATVNLTVIRCPQDTGGGINIFGVLS 228  
Db 383 SNLR----ILGVKSDGFGYQCVAEANEAGNAQTSQAQLIVPKPAIPSS-----VLP 430  
Qy 229 LP 230  
Db 431 AP 432

## RESULT 4

PCT-US94-05277-2  
; Sequence 2, Application PC/TUS9405277  
; GENERAL INFORMATION:  
; APPLICANT: Bruskin, Arthur  
; APPLICANT: Jarosz, David E.  
; APPLICANT: Johnson, Karen  
; APPLICANT: Kinzler, Kenneth W.  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Zdzienicka, James R.  
; TITLE OF INVENTION: Antibodies Specific for DCC Gene Product  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner, Birch, McKie & Beckett  
 STREET: 1001 G Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20001  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/05277  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kagan, Sarah A.  
 REGISTRATION NUMBER: 32,141  
 REFERENCE/DOCKET NUMBER: 01107.42709  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202.508.9100  
 TELEFAX: 202.508.9299  
 TELEX: 197430 BMBB UT  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1447 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US94-05277-2

Query Match 6.2%; Score 128; DB 5; Length 1447;

Best Local Similarity 24.8%; Pred. No. 0.015; 99; Indels 50; Gaps 10;  
 Matches 60; Conservative 33; Mismatches 99;

QY 10 SSGNEV-----LEGQNAVTLKGSQARFNCTVSQGW--KLIMWLSDM 51  
 DB 220 SRTGNEAEVILSDPGLRQLVFLQRPNSVVAIEGKDAVLECCVS-GYPPPSFTWLGRGE 278  
 QY 52 VLSVRPMEPIITNDRTFSQRYDQGNFTSEMIHNVEPDSGNIRC--SLONSRLHGS 109  
 DB 279 VI-----QLRSKYSLLGG--SNLLISNVTDSDSGMYTCVITYKNIENISASA 323  
 QY 110 YLTVOVMGELFIPSVNLVAENEPCEVTLPSHWTRLPDISW-ELGLLVSHSSYFVPEP 168  
 DB 324 ELTVLPVFPFLNHPNSNLVAYESMDIEPECTVS-GKPEVTNWMKNGDVIPSDYFQIVGG 382  
 QY 169 SDLGAVSILALTPOSGNLTLCVATWKLKARKSATVNLTVIRCPQDTGGGINIPGVLS 228  
 DB 383 SNLR----ILGVKSDGEFYQCVAENEAQNAQTSQAQLIVPKPAIPSSS-----VLPS 430  
 QY 229 LP 230  
 DB 431 AP 432

# RESULT 5

US-08-374-834-16  
 Sequence 16, Application US/08374834  
 Patent No. 5656473  
 GENERAL INFORMATION:  
 APPLICANT: Valenzuela, et al.  
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
 STREET: 777 Old Saw Mill River Road  
 CITY: Tarrytown  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10591  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/374,834  
 FILING DATE: 19-JAN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/095,658  
 FILING DATE: 21-JUL-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cobert, Robert J.  
 REGISTRATION NUMBER: 36,108  
 REFERENCE/DOCKET NUMBER: REG 190A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (914) 345-7400  
 TELEFAX: (914) 345-7721  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 869 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-08-374-834-16

Query Match 6.1%; Score 126.5; DB 1; Length 869;

Best Local Similarity 20.9%; Pred. No. 0.01;  
 Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

QY 15 EVIEGPQNAVTLKGSQARFNCTVSQGW--KLIMWLSDMVLSVRPMEPIITNDRTFSQRY 73  
 DB 122 KITRPPINVKIIEGLKAVLPCTTMGNPKPSVSWIKGD-----SPURENSRIAVLE- 171  
 QY 74 DQGNFTSEMIHNVEPDSGNIRCQLONSRLHGSAY-LTVQVMGELFIPSVNLVVAENE 132  
 DB 172 -----SGSLRIHNQVEDAGQVRCVAKNSL--GTAYSKVVKLEVEVFARILRAPESHNV 223  
 QY 133 P-----CEVTLPSHWTRLPDISW-ELGLLVSHSSYFVPEPSDLOSASVILALTQ 183  
 DB 224 TFGSFVTLHCTATGIP-----VPTITWIENGNAVSSGSIQESVKORVIDRSLQLFITKP- 277  
 QY 184 SNTLTCVAT---WKSLLKARKSATVNLTVIRCPQDTGGG-----INIPGVLSLP 230  
 DB 278 --GLYTCIATNKHGEKFEAKAAHISIAENSKPQXDNKGCAQYRGEVCNAVLAADLV 335  
 QY 231 SLGFSL-----PTWGVGLGLAGTMLLTPTCTLTIRCCCCRRRCGCGN----- 273  
 DB 336 FLNTSYADPEEAQELVHTAMNEL-----KVVSPVCRPAEAALLCNHIFQECSPGVVP 388  
 QY 274 ----CCRCRC-----FCCRRKRGFPRIQFKKSEKTKNETETESGNENSGVNSDEQKT 323  
 DB 389 TPIPICREYCLAVKELFCAKE---WLVMEETHRGLYRSEMHLLSVPECSKLPSMHWDP 444  
 QY 324 TDTASLPPKSCSSDPEQRNSSCGPPHQADQRP-----PRPASHFQASFLA 371  
 DB 445 TACARLP-----HLDYKNENLKTFFP--MTSSKPSVDIPNLPSSSSSSSSFSVS 489

# RESULT 6

US-08-644-271-29  
 Sequence 29, Application US/08644271  
 Patent No. 5814478  
 GENERAL INFORMATION:  
 APPLICANT: Valenzuela, et al.  
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS  
 TITLE OF INVENTION: AND LIGANDS  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
 STREET: 777 Old Saw Mill Road  
 CITY: Tarrytown  
 STATE: NY

COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/644,271  
FILING DATE: 10-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 60/008,657  
FILING DATE: 15-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Cobert, Robert J  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 195A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 869 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-644-271-29

Query Match 6.1%; Score 126.5; DB 2; Length 869;  
Best Local Similarity 20.9%; Pred. No. 0.01;  
Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;  
QY 15 EVIEGPQATVLKGSQARENCTVSQGWK-LIMWALSDMVLSVRPMEPIITNDRFTSQRY 73  
DB 122 KITRPINVKIIEGLKAVLPCTMGNPSPSVSWIKGD-----SPLRENSRIAVLE- 171  
QY 74 DOGNGFTSMIHNVPSPDSGNIRCSLQNSRLHGSAY-LTVQVMGELFIPSVNLVVAENE 132  
DB 172 -----SGSLRIHNQVEDAGQYRCVAKNSL--GTAYSKVVKLEVEVFARILRAPESHV 223  
QY 133 P-----CEVTCPLSHWTRLPDISW-ELGLLVSHSSYFVPSPSDLSQSAVSLALTPQ 183  
DB 224 TFGSFVTLHCTAGIP-----VPTIWIENGNAVSSGSIQESVKDVRIDSLQLFITKP- 277  
QY 184 SNGTLTCVAT---WKSLKARKSATVNLTVIRCPQDTGGG-----INIPGVLSLP 230  
DB 278 --GLYTCIATNKHGKRFSTAKAAATISIAEWSKPKQDNKGCAQYRGEVCNAVLAKDALV 335  
QY 231 SLGFSL-----PTWKGVLGLAGTLLPTCTLTIRCCCRRCGQCN----- 273  
DB 336 FLNTSVADPEEAQELLVHTAMNEL-----KVSFVCPRAEALLCNHIFQECSPGVVP 388  
QY 274 ----CCRC--FCCRRKRGFRIQPKKSEKKTNETETESGNENSGYNSDEQKT 323  
DB 389 TPIPCREYCLAVKELFCAKE---WLVMEKTHRGLYRSEMHLLSVPECSKLPMSHMDP 444  
QY 324 TDTASLPKSCSDSPQNSCGPHQADQRP-----PRPASHPQASFNL 371  
DB 445 TACARLP-----HLDYNKENLKTFFP--MTSSKPSVDIENLPSSSSSSSFSVS 489

## RESULT 7

US-09-077-955-33  
Sequence 33, Application US/09077955A  
Patent No. 6413740  
GENERAL INFORMATION:  
APPLICANT: Valenzuela et al., David M.  
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS  
FILE REFERENCE: REG195-B-PCT-US  
CURRENT APPLICATION NUMBER: US/09/077,955A

CURRENT FILING DATE: 1998-09-10  
EARLIER APPLICATION NUMBER: PCT/US96/20696  
EARLIER FILING DATE: 1996-12-13  
EARLIER APPLICATION NUMBER: 08/644,271  
EARLIER FILING DATE: 1996-05-10  
EARLIER APPLICATION NUMBER: 60/008,657  
EARLIER FILING DATE: 1995-12-15  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 33  
LENGTH: 869  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-077-955-33

Query Match 6.1%; Score 126.5; DB 4; Length 869;  
Best Local Similarity 20.9%; Pred. No. 0.01;  
Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;  
QY 15 EVIEGPQATVLKGSQARENCTVSQGWK-LIMWALSDMVLSVRPMEPIITNDRFTSQRY 73  
DB 122 KITRPINVKIIEGLKAVLPCTMGNPSPSVSWIKGD-----SPLRENSRIAVLE- 171  
QY 74 DOGNGFTSMIHNVPSPDSGNIRCSLQNSRLHGSAY-LTVQVMGELFIPSVNLVVAENE 132  
DB 172 -----SGSLRIHNQVEDAGQYRCVAKNSL--GTAYSKVVKLEVEVFARILRAPESHV 223  
QY 133 P-----CEVTCPLSHWTRLPDISW-ELGLLVSHSSYFVPSPSDLSQSAVSLALTPQ 183  
DB 224 TFGSFVTLHCTAGIP-----VPTIWIENGNAVSSGSIQESVKDVRIDSLQLFITKP- 277  
QY 184 SNGTLTCVAT---WKSLKARKSATVNLTVIRCPQDTGGG-----INIPGVLSLP 230  
DB 278 --GLYTCIATNKHGKRFSTAKAAATISIAEWSKPKQDNKGCAQYRGEVCNAVLAKDALV 335  
QY 231 SLGFSL-----PTWKGVLGLAGTLLPTCTLTIRCCCRRCGQCN----- 273  
DB 336 FLNTSVADPEEAQELLVHTAMNEL-----KVSFVCPRAEALLCNHIFQECSPGVVP 388  
QY 274 ----CCRC--FCCRRKRGFRIQPKKSEKKTNETETESGNENSGYNSDEQKT 323  
DB 389 TPIPCREYCLAVKELFCAKE---WLVMEKTHRGLYRSEMHLLSVPECSKLPMSHMDP 444  
QY 324 TDTASLPKSCSDSPQNSCGPHQADQRP-----PRPASHPQASFNL 371  
DB 445 TACARLP-----HLDYNKENLKTFFP--MTSSKPSVDIENLPSSSSSSSFSVS 489

## RESULT 8

US-09-062-365-1  
Sequence 1, Application US/09062365  
Patent No. 6465422  
GENERAL INFORMATION:  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Stern, David  
TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A  
FILE REFERENCE: 55424  
CURRENT APPLICATION NUMBER: US/09/062,365  
CURRENT FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 1  
LENGTH: 332  
TYPE: PRT  
ORGANISM: Human  
US-09-062-365-1

Query Match 6.0%; Score 123.5; DB 4; Length 332;  
Best Local Similarity 24.1%; Pred. No. 0.0049;  
Matches 65; Conservative 33; Mismatches 89; Indels 83; Gaps 14;  
QY 15 EVIEGPQATVLKGSQARENCTVSQGWK-LIMWALSDMVLSVRPMEPIITNDRFTS- 70

Db 103 EIVDSASELTA--GVPNKVGTVCSEGGYPAGTSLNHLGD-----KPLVPNEKGVSV 151  
Qy 71 ----QRYDQGNFT--SMIHNHVPDSGNIR-----CSLQNSRLHGSAYLTQVMGELF 120  
Db 152 KEOTRHPETGLFTLQSLM---VTPARGDPRTFTSCFSPGLPRHRLATAPIQRVW 208  
Qy 121 IP-----SVNLVAENEP-----CEVTCPLPSHWTRLPDLSWELGLVSHSSY 162  
Db 209 EFPLEEVQLVW---EPGGVAPOGTVTLCEVPAQPS-----PQIHWKMD----- 252  
Qy 163 YFVPEPSDLQSAVSIALTPQSGNTLTCVATWKSARKSATVNLTVIRCPDQDGGINI 222  
Db 253 -GVPLPLPSPVLILPEIGPDQGTYSVATHSHGPOESRAVSISIE-PGEG----- 305  
Qy 223 PGLVSSLFSLGSLFTWKGVLGAGTMLL 252  
Db 306 -----FTAGSVGGSGLGTAL 321

## RESULT 9

US-09-651-200-2

; Sequence 2, Application US/09651200

; Patent No. 6429303

; GENERAL INFORMATION:

; APPLICANT: Green et al

; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B

; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and

; TITLE OF INVENTION: Polypeptides Encoded Thereby

; FILE REFERENCE: 15966-562 (CURA-62)

; CURRENT APPLICATION NUMBER: US/09/651,200

; CURRENT FILING DATE: 2000-08-30

; PRIOR APPLICATION NUMBER: 60/152383

; PRIOR FILING DATE: 1999-09-03

; PRIOR APPLICATION NUMBER: 60/172909

; PRIOR FILING DATE: 1999-12-21

; PRIOR APPLICATION NUMBER: 60/183578

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 2

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-651-200-2

Query Match 5.9%; Score 122.5; DB 4; Length 340;  
Best Local Similarity 21.2%; Pred. No. 0.0062;  
Matches 77; Conservative 50; Mismatches 128; Indels 109; Gaps 17;

Qy 10 SSGSNEVIEGPQNAV-LKGSQARFNCTVS--QGWKL-----IMWALSDMVVLVSRPMEPI 62  
Db 48 SPTGAVEVQVPEDPVVALVGTDLHCSFSPPEGSLTQLNLIMQLTDTKQLV----- 100  
Qy 63 ITNDRFTSQRYDQGNF-----TSEMIHNHVPDSGNIRCSLQNSRLHGS 108  
Db 101 ---HSFTEGR-DQGSAYANRTALFPDQLAQNASRLQVRVADEGSFTCFV-SIRDFGS 155  
Qy 109 AYLTVQVMGSLFIPSVNLV-----VAENEPCEVTCLPSHWTRLP--DISWELGL---LVSH 159  
Db 156 AAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDGGVPLTGN 213  
Qy 160 SSVYFVPEPSDLQSAVSIALTPQSGNTLTCVATWKSARKSATVNLTVIRCP---ODT 216  
Db 214 VITSQMANEQGLFDVHSLRVLVGANGTYSC-----LVRNPLVQODA 255  
Qy 217 GGGINIPGVLSLPSLGFSLPTWKGVLGAGTMLLTPTCTLTIRCCCRRCGCCGCC 276  
Db 256 HGSVTTITQDPMTFPEAL-----WTVGLSVCLIALLV----- 288  
Qy 277 RCFCCRRKGGFRIOQKSEKKTETETESGNSGNSDEQKTTDTASLPKSCS 336  
Db 289 ALAFVCRWK-----IKQSCBENAGAEODG-----EGGSKTALQPLKXSDS 331

Qy 337 SDPE 340  
Db 332 KEDD 335

## RESULT 10

US-09-651-200-4

; Sequence 4, Application US/09651200

; Patent No. 6429303

; GENERAL INFORMATION:

; APPLICANT: Green et al

; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B

; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and

; TITLE OF INVENTION: Polypeptides Encoded Thereby

; FILE REFERENCE: 15966-562 (CURA-62)

; CURRENT APPLICATION NUMBER: US/09/651,200

; CURRENT FILING DATE: 2000-08-30

; PRIOR APPLICATION NUMBER: 60/152383

; PRIOR FILING DATE: 1999-09-03

; PRIOR APPLICATION NUMBER: 60/172909

; PRIOR FILING DATE: 1999-12-21

; PRIOR APPLICATION NUMBER: 60/183578

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 4

; LENGTH: 441

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-651-200-4

Query Match 5.9%; Score 122.5; DB 4; Length 441;  
Best Local Similarity 21.2%; Pred. No. 0.0089;  
Matches 77; Conservative 50; Mismatches 128; Indels 109; Gaps 17;

Qy 10 SSGSNEVIEGPQNAV-LKGSQARFNCTVS--QGWKL-----IMWALSDMVVLVSRPMEPI 62  
Db 149 SPTGAVEVQVPEDPVVALVGTDLHCSFSPPEGSLTQLNLIMQLTDTKQLV----- 201  
Qy 63 ITNDRFTSQRYDQGNF-----TSEMIHNHVPDSGNIRCSLQNSRLHGS 108  
Db 202 ---HSFTEGR-DQGSAYANRTALFPDQLAQNASRLQVRVADEGSFTCFV-SIRDFGS 256  
Qy 109 AYLTVQVMGSLFIPSVNLV-----VAENEPCEVTCLPSHWTRLP--DISWELGL---LVSH 159  
Db 257 AAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDGGVPLTGN 314  
Qy 160 SSVYFVPEPSDLQSAVSIALTPQSGNTLTCVATWKSARKSATVNLTVIRCP---ODT 216  
Db 315 VITSQMANEQGLFDVHSLRVLVGANGTYSC-----LVRNPLVQODA 356  
Qy 217 GGGINIPGVLSLPSLGFSLPTWKGVLGAGTMLLTPTCTLTIRCCCRRCGCCGCC 276  
Db 357 HGSVTTITQDPMTFPEAL-----WTVGLSVCLIALLV----- 389  
Qy 277 RCFCCRRKGGFRIOQKSEKKTETETESGNSGNSDEQKTTDTASLPKSCS 336  
Db 390 ALAFVCRWK-----IKQSCBENAGAEODG-----EGGSKTALQPLKXSDS 432  
Qy 337 SDPE 340  
Db 433 KEDD 436

## RESULT 11

US-09-651-200-6

; Sequence 6, Application US/09651200

; Patent No. 6429303

; GENERAL INFORMATION:

; APPLICANT: Green et al

; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B

; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and

; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 15966-562 (CURA-62)  
; CURRENT APPLICATION NUMBER: US/09/651,200  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/152383  
; PRIOR FILING DATE: 1999-09-03  
; PRIOR APPLICATION NUMBER: 60/172909  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/183578  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-651-200-6

Query Match 5.9%; Score 121.5; DB 4; Length 534;  
Best Local Similarity 21.2%; Pred. No. 0.014;  
Matches 77; Conservative 49; Mismatches 129; Indels 109; Gaps 17;  
QY 10 SSGNEVIEGPNATV-LKGSQARNCTVS--QGWKL-----IMWALSDMVLSVRPMEPI 62  
Db 242 SPTGAVEVQVPEDPVVALVGTDTLRCSFSPGSLAQLNLIWQLTDTKQLV----- 294  
QY 63 ITNDRPTSQRVDQGNF-----TSEMIHNVPSPSGNIRCSLQNSRLHGS 108  
Db 295 ---HSFTEGR-DQGSAYANRTALFPDLLAQGNASRLQVRVADGSGFTCFV-SIRDFGS 349  
QY 109 AVLTQVMGELFIPSNLV-----VAENPECEVTCLPSPHWTRLP--DISWELGL---LVSH 159  
Db 350 AAVSLQVAAPYKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDQGVPLTGN 407  
QY 160 SSYFVPEPSDLQSAVSIALTPQSNGLTLCVATWKSILKARKSATVNLTVIRCP---QDT 216  
Db 408 VTTQSMANEQGLFDVHSLRVVLGANGTYSC-----LVNPNVLQODA 449  
QY 217 GGGINIPGVLSLPSLPGSLPTWKGVLGLAGTMLLTCTLTIRCCCRRCGCCNCCC 276  
Db 450 HGSVITITGQMTFFPEAL-----WVTVGLSVCLIALLV----- 482  
QY 277 RCCFCRRKRGRIQFKKSEKTKETETESGNENSGYNSDSQKTTDTASLPKSCES 336  
Db 483 ALAFVCWRK-----IKQSCSEENAGAEQDQ-----EGGSKTALQPLKHSDS 525  
QY 337 SDPE 340  
Db 526 KEDD 529

RESULT 12  
US-09-651-200-24  
; Sequence 24, Application US/09651200  
; Patent No. 6429303  
; GENERAL INFORMATION:  
; APPLICANT: Green et al  
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B  
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 15966-562 (CURA-62)  
; CURRENT APPLICATION NUMBER: US/09/651,200  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/152383  
; PRIOR FILING DATE: 1999-09-03  
; PRIOR APPLICATION NUMBER: 60/172909  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/183578  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 534

; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Sequence  
; OTHER INFORMATION: mz5020.protein from Figure 4.  
US-09-651-200-24  
Query Match 5.9%; Score 121.5; DB 4; Length 534;  
Best Local Similarity 21.2%; Pred. No. 0.014;  
Matches 77; Conservative 49; Mismatches 129; Indels 109; Gaps 17;  
QY 10 SSGNEVIEGPNATV-LKGSQARNCTVS--QGWKL-----IMWALSDMVLSVRPMEPI 62  
Db 242 SPTGAVEVQVPEDPVVALVGTDTLRCSFSPGSLAQLNLIWQLTDTKQLV----- 294  
QY 63 ITNDRPTSQRVDQGNF-----TSEMIHNVPSPSGNIRCSLQNSRLHGS 108  
Db 295 ---HSFTEGR-DQGSAYANRTALFPDLLAQGNASRLQVRVADGSGFTCFV-SIRDFGS 349  
QY 109 AVLTQVMGELFIPSNLV-----VAENPECEVTCLPSPHWTRLP--DISWELGL---LVSH 159  
Db 350 AAVSLQVAAPYKPSMTLEPNKDLRPGDVTITC--SSYRGYPEAEVFWQDQGVPLTGN 407  
QY 160 SSYFVPEPSDLQSAVSIALTPQSNGLTLCVATWKSILKARKSATVNLTVIRCP---QDT 216  
Db 408 VTTQSMANEQGLFDVHSLRVVLGANGTYSC-----LVNPNVLQODA 449  
QY 217 GGGINIPGVLSLPSLPGSLPTWKGVLGLAGTMLLTCTLTIRCCCRRCGCCNCCC 276  
Db 450 HGSVITITGQMTFFPEAL-----WVTVGLSVCLIALLV----- 482  
QY 277 RCCFCRRKRGRIQFKKSEKTKETETESGNENSGYNSDSQKTTDTASLPKSCES 336  
Db 483 ALAFVCWRK-----IKQSCSEENAGAEQDQ-----EGGSKTALQPLKHSDS 525  
QY 337 SDPE 340  
Db 526 KEDD 529

RESULT 13  
US-08-633-148-4  
; Sequence 4, Application US/08633148  
; Patent No. 5864018  
; GENERAL INFORMATION:  
; APPLICANT: MORSE, MICHAEL J.  
; APPLICANT: NAGASHIMA, NARIKO  
; APPLICANT: HOLLANDER, DORIS A.  
; TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: END-PRODUCT RECEPTOR PEPTIDES AND USES THEREFOR  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP  
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: U.S.A.  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/633,148  
; FILING DATE: 16-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURPHY ESQ., MATTHEW B.  
; REGISTRATION NUMBER: 39,787  
; REFERENCE/DOCKET NUMBER: 014618-00560005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400

```

; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-633-148-4

Query Match          5.8%; Score 118.5; DB 2; Length 318;
Best Local Similarity 24.0%; Pred. No. 0.013;
Matches 64; Conservative 33; Mismatches 88; Indels 83; Gaps 14;

Qy 15 EVIEGPONATVILKGSQARFNCTVSQ---WKLIMWALSDMVLSVRPMEPIITNDRFTS- 70
Db 103 EIVDSASELTA--GVPNKVGTCVSEGSYPAGTILSWHLDG-----KPLVNEKGVSV 151
Qy 71 ----QRYDQGNFT--SEMIHNVPSDSGNIR----CSLQNSRLHGSAYLTVQVMGELF 120
Db 152 KEQTRRHPETGLTLOSELM---VTPARGDPRPTSCSPCLPHEHRLARTAPIQPRVM 208
Qy 121 IP-----SVNLVVAENP-----CVTCLPSHWTLPDLSWELGLLVSHSSY 162
Db 209 EPVPLEEVQLW---EPEGGAAPGGTVTLTCEVPAQPS-----POIHWMKD----- 252
Qy 163 YFVPEPSDLQSAVSIILALTPQSNGLTLCVATKSLKARKSATVNLVIRCPQDTGGGINI 222
Db 253 -GVPLPLPPSPVLLPEIGPDQGTYSVATHSHGQPSRAVSIISII-PGEEG----- 305
Qy 223 FGVLSLPSLGSFSLPTWKGVLGLAGT 249
Db 306 -----PTAGSVGGSGGLGT 318

```

## RESULT 14

```

US-08-633-148-2
; Sequence 2, Application US/08633148
; Patent No. 5864018
; GENERAL INFORMATION:
; APPLICANT: MORSE, MICHAEL J.
; APPLICANT: NAGASHIMA, MARIKO
; APPLICANT: HOLLANDER, DORIS A.
; TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND & TOWNSEND & CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,148
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY ESQ., MATTHEW B.
; REGISTRATION NUMBER: 39,787
; REFERENCE/DOCKET NUMBER: 014618-00560005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-633-148-2

Query Match          5.8%; Score 118.5; DB 2; Length 340;
Best Local Similarity 24.0%; Pred. No. 0.014;
Matches 64; Conservative 32; Mismatches 88; Indels 83; Gaps 14;

Qy 15 EVIEGPONATVILKGSQARFNCTVSQ---WKLIMWALSDMVLSVRPMEPIITNDRFTS- 70
Db 125 EIVDSASELTA--GVPNKVGTCVSEGSYPAGTILSWHLDG-----KPLVNEKGVSV 173
Qy 71 ----QRYDQGNFT--SEMIHNVPSDSGNIR----CSLQNSRLHGSAYLTVQVMGELF 120
Db 174 KEQTRRHPETGLTLOSELM---VTPARGDPRPTSCSPCLPHEHRLARTAPIQPRVM 230
Qy 121 IP-----SVNLVVAENP-----CVTCLPSHWTLPDLSWELGLLVSHSSY 162
Db 231 EPVPLEEVQLW---EPEGGAAPGGTVTLTCEVPAQPS-----POIHWMKD----- 274
Qy 163 YFVPEPSDLQSAVSIILALTPQSNGLTLCVATKSLKARKSATVNLVIRCPQDTGGGINI 222
Db 275 -GVPLPLPPSPVLLPEIGPDQGTYSVATHSHGQPSRAVSIISII-PGEEG----- 327
Qy 223 FGVLSLPSLGSFSLPTWKGVLGLAGT 249
Db 328 -----PTAGSVGGSGGLGT 340

```

## RESULT 15

```

US-09-651-200-20
; Sequence 20, Application US/09651200
; Patent No. 6429303
; GENERAL INFORMATION:
; APPLICANT: Green et al
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-562 (CURA-62)
; CURRENT APPLICATION NUMBER: US/09/651,200
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152383
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/172909
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/183578
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 20
; LENGTH: 325
; TYPE: PRT
; ORGANISM: sus sp.
; US-09-651-200-20

Query Match          5.7%; Score 117; DB 4; Length 325;
Best Local Similarity 22.9%; Pred. No. 0.018;
Matches 88; Conservative 39; Mismatches 126; Indels 132; Gaps 22;

Qy 3 LVIFLHSGSGNEVIEGPNATVILKGSQARFN-----C--TVSQGWL-----IMWALSD 50
Db 10 MVLLLSGAAS-----L-K-SQAYFNETGELPCHFTNSQLSLDELVIFWQDQD 55
Qy 51 MVVL--SVRPME--PIITNDRFTSQ--RYDQGNFTTSEMIHNVPSDSGNIRCSLQNSRLH 106
Db 56 NLVLYELRGQEKPHNVNRYKMGRTSPDQA---TWILRLHNVQIKKSGVQCQFIHKGPH 112
Qy 107 GSAYL-----TVQVMGELFIPSVNLVV--AENPEFCEVTCPSHWTLPDLSWELGLLVSH 159
Db 113 GLVPIHQWSSDLSLLANFSQPEINLLTNHTNSVINLTC-----151
Qy 160 SSVYFVPEPSDLQSAVSIILALTPQSNGLTLCVATKSLKARKSATVNLTV---IRCPQDT 216

```

Search completed: November 13, 2003, 03:35:11  
Job time : 15.2478 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:34:36 ; Search time 26.4818 Seconds  
(without alignments)  
2477.484 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059

Sequence: 1 MGLVIFLHGGSGNEVEIGP.....HPQASFNLASPEKVSNTTW 382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	6.9	405	US-08-755-235-4	Sequence 4, Appli
2	140	6.8	633	US-10-180-410-26	Sequence 26, Appl
3	138.5	6.7	2473	US-10-184-644-559	Sequence 559, App
4	138.5	6.7	2473	US-10-184-634-559	Sequence 559, App
5	138	6.7	592	US-10-180-410-2	Sequence 2, Appli
6	136	6.6	594	US-10-180-410-12	Sequence 12, Appl
7	136	6.6	708	US-10-180-410-8	Sequence 8, Appli
8	136	6.6	708	US-10-199-672-584	Sequence 584, App
9	136	6.6	708	US-10-187-749-584	Sequence 584, App
10	136	6.6	708	US-10-194-457-584	Sequence 584, App
11	136	6.6	708	US-10-184-642-584	Sequence 584, App
12	136	6.6	708	US-10-196-747-584	Sequence 584, App
13	136	6.6	708	US-10-173-689-584	Sequence 584, App
14	136	6.6	708	US-10-173-690-584	Sequence 584, App
15	136	6.6	708	US-10-173-691-584	Sequence 584, App

16	136	6.6	708	12	US-10-173-692-584	Sequence 584, App
17	136	6.6	708	12	US-10-173-694-584	Sequence 584, App
18	136	6.6	708	12	US-10-173-698-584	Sequence 584, App
19	136	6.6	708	12	US-10-173-699-584	Sequence 584, App
20	136	6.6	708	12	US-10-173-707-584	Sequence 584, App
21	136	6.6	708	12	US-10-174-569-584	Sequence 584, App
22	136	6.6	708	12	US-10-174-583-584	Sequence 584, App
23	136	6.6	708	12	US-10-174-587-584	Sequence 584, App
24	136	6.6	708	12	US-10-174-589-584	Sequence 584, App
25	136	6.6	708	12	US-10-175-591-584	Sequence 584, App
26	136	6.6	708	12	US-10-175-736-584	Sequence 584, App
27	136	6.6	708	12	US-10-175-743-584	Sequence 584, App
28	136	6.6	708	12	US-10-175-744-584	Sequence 584, App
29	136	6.6	708	12	US-10-175-745-584	Sequence 584, App
30	136	6.6	708	12	US-10-175-748-584	Sequence 584, App
31	136	6.6	708	12	US-10-175-751-584	Sequence 584, App
32	136	6.6	708	12	US-10-175-754-584	Sequence 584, App
33	136	6.6	708	12	US-10-176-480-584	Sequence 584, App
34	136	6.6	708	12	US-10-176-489-584	Sequence 584, App
35	136	6.6	708	12	US-10-176-754-584	Sequence 584, App
36	136	6.6	708	12	US-10-176-755-584	Sequence 584, App
37	136	6.6	708	12	US-10-176-759-584	Sequence 584, App
38	136	6.6	708	12	US-10-176-920-584	Sequence 584, App
39	136	6.6	708	12	US-10-176-922-584	Sequence 584, App
40	136	6.6	708	12	US-10-176-924-584	Sequence 584, App
41	136	6.6	708	12	US-10-179-508-584	Sequence 584, App
42	136	6.6	708	12	US-10-179-508-584	Sequence 584, App
43	136	6.6	708	12	US-10-179-512-584	Sequence 584, App
44	136	6.6	708	12	US-10-179-515-584	Sequence 584, App
45	136	6.6	708	12	US-10-173-702-584	Sequence 584, App

## ALIGNMENTS

### RESULT 1

US-08-755-235-4  
; Sequence 4, Application US/08755235  
; Publication No. US20030059423A1  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Wu, Jun  
; TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES  
; FILE REFERENCE: 0575/50159  
; CURRENT APPLICATION NUMBER: US/08/755.235  
; CURRENT FILING DATE: 1996-11-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Human  
US-08-755-235-4

Query Match	6.9%;	Score 142;	DB 8;	Length 405;
Best Local Similarity	22.8%;	Pred. No. 0.00067;		
Matches	86;	Conservative 40;	Mismatches 113;	Indels 138; Gaps 19;
QY	15	EVIEGPONATVILKGSQARFNCVTSQG---	WKLIMWALSDMVLSVRPMEPIITNDRTFS-	70
Db	125	EIVDSASELTA--GVNPKVGTCSSEGSYPAGTLSWHLDG-----	KPLVFNKEGVSV	173
QY	71	-----QRYDQGNFT--SEMIHNYEPPSDSGNIR-----	CSLQNSKLHSGASVLTQVMGELF	120
Db	174	KEQTRRHETGLFTLQSELM---VTPARGGDPRTFCSPGIPRHRALRTAPIOPRVW	230	
QY	121	IP-----SYNLVAENEP-----	CEVTCPLPSHWTRLPDISWELGLVSHSSY	162
Db	231	EVVPLEEVQLVV---EPEGGAVAGGTVTLTCEVPAQPS-----	PQIHWKMD-----	274
QY	163	YFVFPSPDLQSAVSIALTPOSGNLTCTCATWTKSLKARKS-ATVNLTVIRCPQDTGGGIN	221	

```
Db 275 -GVPLPSPVLLIPETGPQDQGYSCVATHSHGPOESRAVVISIIE-PGEEG----- 328
Qy 222 IPGVLSLPSLGSFTWGVGLGLAGTMLTF-----PTLTIRCCCRCCGNC 274
Db 329 -----PTAGSVGSGGLTALALGILGGLTAALLIGVILWQR----- 367
Qy 275 CCECCFCCRRKGRGFIQFKSEKKT--NKETETESGNEGSGYNSDQKQTTDTASLPK 332
Db 369 -----QRRGERKAPENQEESEAEIN----- 390
Qy 333 SCESSDPEQRNSSCGPP 349
Db 391 --QSEEPAGESSGTGP 405

RESULT 2
US-10-180-410-26
; Sequence 26, Application US/10180410
; Publication No. US20030148382A1
; GENERAL INFORMATION:
; APPLICANT: SUN, CHAO
; APPLICANT: CARULLI, JOHN P.
; APPLICANT: LORASHIN, ALEXANDER V.
; APPLICANT: KILBURN, DANIEL R.
; TITLE OF INVENTION: PANCAM NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A097 CIP
; CURRENT APPLICATION NUMBER: US/10/180,410
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: PCT/US01/19904
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,611
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 26
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-180-410-26

Query Match 6.8%; Score 140; DB 12; Length 633;
Best Local Similarity 24.2%; Pred. No. 0.0018;
Matches 56; Conservative 40; Mismatches 105; Indels 30; Gaps 10;

Qy 3 LVIFHSGSGNEVEVGPNATVILKSGQARFNCTVSQGWKLIMWALSDMVLSVRPMEPI 62
Db 13 LFCFRGRGSPHFLOQPEDVLVLLGEEARLPCALGAYWGLVQWTKSGLALGGQR----- 67
Qy 63 ITNDRFTRQYDQGNFTS---EMITHNVEPDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLFGWRYWISGNAANGQHDHLPVLEDEASQYEQATQAGLRSPALQHLVLPPE 124
Qy 114 --QVMGELFIPSNLVVAENPECVTCLPSHWTR-LPDISW-ELGLLVSHSSY--FVPE 167
Db 125 APQVLGG---PSVSLVA--GVFANLTCRSRGDARPTPELLWFRDGLDGFATHTLKE 179
Qy 168 --PSDLOSQVSTLALPQSNGLTTCVATWKSLSKARSAATVNLTVIRCPQDT 216
Db 180 GTPGCVSESTLTPFSDHDDGATLVCRARSQALPTGRDRTAILSLQYPPFVT 230

RESULT 3
US-10-184-644-559
; Sequence 559, Application US/10184644
; Publication No. US2003004930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
```

```
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 559
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-559

Query Match 6.7%; Score 138.5; DB 15; Length 2473;
Best Local Similarity 29.5%; Pred. No. 0.016;
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

Qy 179 ALTPQSNGLTTCVATWKSLSKARSAATVNLTVIRCPQDTGGGINIPGVLSLSLGSLSLT 238
Db 2274 AATTGAAGTTTCAATTAATAATTAATGTTTCC----- 2307
Qy 239 WGVVGLGLAGTMLLTPTCTLTIRCCCR-RRCCGCMCCCRCCFCC 282
Db 2308 -----ATTCTCATGCCGCCACCCCGCCGCCGCCACCCACC 2344

RESULT 4
US-10-184-634-559
; Sequence 559, Application US/10184634
; Publication No. US2003006868A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 559
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-559

Query Match 6.7%; Score 138.5; DB 15; Length 2473;
Best Local Similarity 29.5%; Pred. No. 0.016;
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

Qy 179 ALTPQSNGLTTCVATWKSLSKARSAATVNLTVIRCPQDTGGGINIPGVLSLSLGSLSLT 238
Db 2274 AATTGAAGTTTCAATTAATAATTAATGTTTCC----- 2307
Qy 239 WGVVGLGLAGTMLLTPTCTLTIRCCCR-RRCCGCMCCCRCCFCC 282
Db 2308 -----ATTCTCATGCCGCCACCCCGCCGCCGCCACCCACC 2344

RESULT 5
```

US-10-180-410-2  
 ; Sequence 2, Application US/10180410  
 ; Publication No. US20030148382A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUN, CHAO  
 ; APPLICANT: CARULLI, JOHN P.  
 ; APPLICANT: LUKASHIN, ALEXANDER V.  
 ; APPLICANT: KILBURN, DANIEL R.  
 ; TITLE OF INVENTION: PANCAM NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: A097 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/180,410  
 ; PRIOR FILING DATE: 2002-06-24  
 ; PRIOR APPLICATION NUMBER: PCT/US01/19904  
 ; PRIOR FILING DATE: 2001-06-22  
 ; PRIOR APPLICATION NUMBER: 60/213,611  
 ; PRIOR FILING DATE: 2000-06-22  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 592  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-180-410-2

Query Match 6.7%; Score 138; DB 12; Length 592;  
 Best Local Similarity 23.8%; Pred. No. 0.0025;  
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;  
 QY 3 LVIFLHSGSGNEVIEGPONATVLKGSQARFNCTVSGQWKLIMWALSDMVLSVRPMEPI 62  
 DB 10 LFCFRGAGSPHFLQPPEDLVLLGEEARLPCALGAYWGLVQWTKSGLALGGQR----- 64  
 QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113  
 DB 65 --DLPQWSRYWISGNAANGQHDHIRPVELEDEASVECOATQAGLRSPAQHLVLPPE 121  
 QY 114 --QVMGELFIPSNLVVAENEPCEVTLCPHSHWTR-LPDISW-ELGLLVSHSSYV--FVPE 167  
 DB 122 APQVLGG---PSVSLVA--GVPAULTCRSRGDARPTPELLWFRDGVLLDGTTFHQTLIKE 176  
 QY 168 --PSDLOSASVILALTPQSNGLTTCVATWKSLSKARKSATVNLTVIRCPQDT 216  
 DB 177 GTPGSVESTLTLPFFSHDDGATFVCRARSQALPTGRDGTATILSLQYPPPEVT 227

RESULT 6  
 US-10-180-410-12  
 ; Sequence 12, Application US/10180410  
 ; Publication No. US20030148382A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUN, CHAO  
 ; APPLICANT: CARULLI, JOHN P.  
 ; APPLICANT: LUKASHIN, ALEXANDER V.  
 ; APPLICANT: KILBURN, DANIEL R.  
 ; TITLE OF INVENTION: PANCAM NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: A097 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/180,410  
 ; PRIOR FILING DATE: 2002-06-24  
 ; PRIOR APPLICATION NUMBER: PCT/US01/19904  
 ; PRIOR FILING DATE: 2001-06-22  
 ; PRIOR APPLICATION NUMBER: 60/213,611  
 ; PRIOR FILING DATE: 2000-06-22  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 12  
 ; LENGTH: 594  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-180-410-12

Query Match 6.6%; Score 136; DB 12; Length 594;  
 Best Local Similarity 23.8%; Pred. No. 0.0037;  
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVLKGSQARFNCTVSGQWKLIMWALSDMVLSVRPMEPI 62  
 DB 10 LFCFRGAGSPHFLQPPEDLVLLGEEARLPCALGAYWGLVQWTKSGLALGGQR----- 64  
 QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113  
 DB 65 --DLPQWSRYWISGNAANGQHDHIRPVELEDEASVECOATQAGLRSPAQHLVLPPE 121  
 QY 114 --QVMGELFIPSNLVVAENEPCEVTLCPHSHWTR-LPDISW-ELGLLVSHSSYV--FVPE 167  
 DB 122 APQVLGG---PSVSLVA--GVPAULTCRSRGDARPTPELLWFRDGVLLDGTTFHQTLIKE 176  
 QY 168 --PSDLOSASVILALTPQSNGLTTCVATWKSLSKARKSATVNLTVIRCPQDT 216  
 DB 177 GTPGSVESTLTLPFFSHDDGATFVCRARSQALPTGRDGTATILSLQYPPPEVT 227

RESULT 7  
 US-10-180-410-8  
 ; Sequence 8, Application US/10180410  
 ; Publication No. US20030148382A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUN, CHAO  
 ; APPLICANT: CARULLI, JOHN P.  
 ; APPLICANT: LUKASHIN, ALEXANDER V.  
 ; APPLICANT: KILBURN, DANIEL R.  
 ; TITLE OF INVENTION: PANCAM NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: A097 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/180,410  
 ; CURRENT FILING DATE: 2002-06-24  
 ; PRIOR APPLICATION NUMBER: PCT/US01/19904  
 ; PRIOR FILING DATE: 2001-06-22  
 ; PRIOR APPLICATION NUMBER: 60/213,611  
 ; PRIOR FILING DATE: 2000-06-22  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 708  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-180-410-8

Query Match 6.6%; Score 136; DB 12; Length 708;  
 Best Local Similarity 23.8%; Pred. No. 0.0047;  
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVLKGSQARFNCTVSGQWKLIMWALSDMVLSVRPMEPI 62  
 DB 13 LFCFRGAGSPHFLQPPEDLVLLGEEARLPCALGAYWGLVQWTKSGLALGGQR----- 67  
 QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113  
 DB 68 --DLPQWSRYWISGNAANGQHDHIRPVELEDEASVECOATQAGLRSPAQHLVLPPE 124  
 QY 114 --QVMGELFIPSNLVVAENEPCEVTLCPHSHWTR-LPDISW-ELGLLVSHSSYV--FVPE 167  
 DB 125 APQVLGG---PSVSLVA--GVPAULTCRSRGDARPTPELLWFRDGVLLDGTTFHQTLIKE 179  
 QY 168 --PSDLOSASVILALTPQSNGLTTCVATWKSLSKARKSATVNLTVIRCPQDT 216  
 DB 180 GTPGSVESTLTLPFFSHDDGATFVCRARSQALPTGRDGTATILSLQYPPPEVT 230

RESULT 8  
 US-10-199-672-584  
 ; Sequence 584, Application US/10199672  
 ; Publication No. US20030148442A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC1
; CURRENT APPLICATION NUMBER: US/10/199,672
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-672-584

```

Query Match 6.6%; Score 136; DB 12; Length 708;

Best Local Similarity 23.8%; Pred. No. 0.0047;

Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

```

QY 3 LVIFLHGSGNEVIGPQNAVILKGSQARFNCTVSGQKLIWALSDMVLSVRPMEPI 62
Db 13 LFCFRGRGSPHFLOQPEDLVLLGGEARLPCALGAYWGLVQWTKSGLAGGQR----- 67
QY 63 ITNDRFTSQRYDQGGNFTS---EMIHNVPEPSDGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPQWSRYWISGNAANGQHDHLPVLEDEASVECAQATQAGLRSPQAQLHVLVPPE 124
QY 114 --QVMGELFIPSNLVVAENEPCEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167
Db 125 APQVLGG---PSVSLVA--GVPAULTCRSRGDARPTPELLNFRDGVLLDGAFFHOTLLKE 179
QY 168 --PSDLQSAVSIALTPQSGNGLTCTVATWKSILKARKSATVNLTVIRCPQDT 216
Db 180 GTPGVSVESTLTLPFFSHDDGATFVCRARSQALPTGRTDITLSLQYPPPEVT 230

```

RESULT 9

US-10-187-749-584

; Sequence 584, Application US/10187749

; Publication No. US20030153036A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

```

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC1
; CURRENT APPLICATION NUMBER: US/10/187,749
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-749-584

```

Query Match 6.6%; Score 136; DB 12; Length 708;

Best Local Similarity 23.8%; Pred. No. 0.0047;

Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

```

QY 3 LVIFLHGSGNEVIGPQNAVILKGSQARFNCTVSGQKLIWALSDMVLSVRPMEPI 62
Db 13 LFCFRGRGSPHFLOQPEDLVLLGGEARLPCALGAYWGLVQWTKSGLAGGQR----- 67
QY 63 ITNDRFTSQRYDQGGNFTS---EMIHNVPEPSDGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPQWSRYWISGNAANGQHDHLPVLEDEASVECAQATQAGLRSPQAQLHVLVPPE 124
QY 114 --QVMGELFIPSNLVVAENEPCEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167
Db 125 APQVLGG---PSVSLVA--GVPAULTCRSRGDARPTPELLNFRDGVLLDGAFFHOTLLKE 179
QY 168 --PSDLQSAVSIALTPQSGNGLTCTVATWKSILKARKSATVNLTVIRCPQDT 216
Db 180 GTPGVSVESTLTLPFFSHDDGATFVCRARSQALPTGRTDITLSLQYPPPEVT 230

```

RESULT 10

US-10-194-457-584

; Sequence 584, Application US/10194457

; Publication No. US20030153037A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

```
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C296
; CURRENT APPLICATION NUMBER: US/10/194,457
; CURRENT FILING DATE: 2002-07-11
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/052586
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-457-584

Query Match      6.6%; Score 136; DB 12; Length 708;
Best Local Similarity 23.8%; Pred. No. 0.0047;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVLKSGQARFNCTVSGQWKLIMWALSDMVLSVRPMEPI 62
Db 13 LFCFRGRAGSPHFLOQPEDLVLLGGEARLPALGAYWGLVQWTKSGLAGGQR----- 67
QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVEPPSDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPGWSRYWISGNAANGQHDHLPVPELEDEASVECOATQAGLRSPAQHLVLPPE 124
QY 114 --QWMBELFIPSNLVVAENPEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167
Db 125 APQVLGG---PSVSLVA--GVPAHLTCRSGDARPTPELLWFRDGVLLDGFATPHQTLKE 179
QY 168 --PSDLQASVSIILATPQSNGLTTCVATKSKARKSATVNLTVIRCPDPT 216
Db 180 GTPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTAITSLSQYPPPT 230

RESULT 11
US-10-184-642-584
; Sequence 584, Application US/10184642
; Publication No. US20030157635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C296
; CURRENT APPLICATION NUMBER: US/10/194,457
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-457-584

Query Match      6.6%; Score 136; DB 12; Length 708;
Best Local Similarity 23.8%; Pred. No. 0.0047;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVLKSGQARFNCTVSGQWKLIMWALSDMVLSVRPMEPI 62
Db 13 LFCFRGRAGSPHFLOQPEDLVLLGGEARLPALGAYWGLVQWTKSGLAGGQR----- 67
QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVEPPSDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPGWSRYWISGNAANGQHDHLPVPELEDEASVECOATQAGLRSPAQHLVLPPE 124
QY 114 --QWMBELFIPSNLVVAENPEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167
Db 125 APQVLGG---PSVSLVA--GVPAHLTCRSGDARPTPELLWFRDGVLLDGFATPHQTLKE 179
QY 168 --PSDLQASVSIILATPQSNGLTTCVATKSKARKSATVNLTVIRCPDPT 216
Db 180 GTPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTAITSLSQYPPPT 230

RESULT 12
US-10-196-747-584
; Sequence 584, Application US/10196747
; Publication No. US20030162250A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C346
; CURRENT APPLICATION NUMBER: US/10/196,747
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-747-584

Query Match      6.6%; Score 136; DB 12; Length 708;
Best Local Similarity 23.8%; Pred. No. 0.0047;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVLKSGQARFNCTVSGQWKLIMWALSDMVLSVRPMEPI 62
Db 13 LFCFRGRAGSPHFLOQPEDLVLLGGEARLPALGAYWGLVQWTKSGLAGGQR----- 67
QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVEPPSDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPGWSRYWISGNAANGQHDHLPVPELEDEASVECOATQAGLRSPAQHLVLPPE 124
```

QY 114 --QVMGELFIPSNVLVAENEPCEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167  
Db 125 APQVLGG---PSVSLVA--GVPAULTCRSGDARPTPELLWFRDGVLLDGGATFHQTLLKE 179  
QY 168 --PSDLQSAVSIALTPQSNGLTLCVATWKSLSKARKSATVNLTVIRCPQDT 216  
Db 180 GTPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTRAITLSLOYPPPEVT 230

## RESULT 13

US-10-173-689-584  
; Sequence 584, Application US/10173689  
; Publication No. US20030166104A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C10

; CURRENT APPLICATION NUMBER: US/10/173,689

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 584

; LENGTH: 708

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-689-584

Query Match 6.6%; Score 136; DB 12; Length 708;  
Best Local Similarity 23.8%; Pred. No. 0.0047;  
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHGGSGNEVIEGPONATVLKGSQARFNCTVSQGWKLIMWALS DMVLSVRPMEPI 62  
Db 13 LFCFRGRAGSPHFLQOPEDLVLLGGEARLPCALGAYWGLVQWTKSGLALGGQR----- 67  
QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113  
Db 68 ---DLPGWSRYWISGNAANGQHDHIRPVELEDEASEYECQATQAGLSRPAQLHLVLPPE 124  
QY 114 --QVMGELFIPSNVLVAENEPCEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167  
Db 125 APQVLGG---PSVSLVA--GVPAULTCRSGDARPTPELLWFRDGVLLDGGATFHQTLLKE 179  
QY 168 --PSDLQSAVSIALTPQSNGLTLCVATWKSLSKARKSATVNLTVIRCPQDT 216  
Db 180 GTPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTRAITLSLOYPPPEVT 230

## RESULT 14

US-10-173-690-584  
; Sequence 584, Application US/10173690  
; Publication No. US20030166105A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C9  
; CURRENT APPLICATION NUMBER: US/10/173,690  
; CURRENT FILING DATE: 2002-06-17  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 584  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-173-690-584

Query Match 6.6%; Score 136; DB 12; Length 708;  
Best Local Similarity 23.8%; Pred. No. 0.0047;  
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHGGSGNEVIEGPONATVLKGSQARFNCTVSQGWKLIMWALS DMVLSVRPMEPI 62  
Db 13 LFCFRGRAGSPHFLQOPEDLVLLGGEARLPCALGAYWGLVQWTKSGLALGGQR----- 67  
QY 63 ITNDRFTSQRYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113  
Db 68 ---DLPGWSRYWISGNAANGQHDHIRPVELEDEASEYECQATQAGLSRPAQLHLVLPPE 124  
QY 114 --QVMGELFIPSNVLVAENEPCEVTCPLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167  
Db 125 APQVLGG---PSVSLVA--GVPAULTCRSGDARPTPELLWFRDGVLLDGGATFHQTLLKE 179  
QY 168 --PSDLQSAVSIALTPQSNGLTLCVATWKSLSKARKSATVNLTVIRCPQDT 216  
Db 180 GTPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTRAITLSLOYPPPEVT 230

## RESULT 15

US-10-173-691-584  
; Sequence 584, Application US/10173691  
; Publication No. US20030166106A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C13

; CURRENT APPLICATION NUMBER: US/10/173,691

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 584

; LENGTH: 708

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-691-584

Query Match 6.6%; Score 136; DB 12; Length 708;  
Best Local Similarity 23.8%; Pred. No. 0.0047;  
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHGGSGNEVIEGPONATVLKGSQARFNCTVSQGWKLIMWALS DMVLSVRPMEPI 62  
Db 13 LFCFRGRAGSPHFLQOPEDLVLLGGEARLPCALGAYWGLVQWTKSGLALGGQR----- 67

```

QY 63 ITNDRFTSQRYDOGNFTS---EMIIHNVPESDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPQWSRYWISGNAANGQHDLHIRPVELEDEASVECOATQAGLSRPAQLHVLVPE 124
QY 114 --QVMGELFIPSNLVVAENEPCEVTCLPSHWTR-LPDISW-ELGLLVSHSSYY--FVPE 167
Db 125 APQVLGG---PSVSLVA--GVSPANLTCRSRGDARPTPELIWFRDGVLLDGAIFHOTLLKE 179
QY 168 --PSDLQSAVSILALTPQSNGLTLCVATWKSARKSATVNLTVIRCPDUT 216
Db 180 GTPGSVESILTITPFSSHDDGATFVCRARSOALPTGRDTAILSLQYPPVT 230

```

Search completed: November 13, 2003, 03:46:13  
Job time : 29.4818 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:40:51 ; Search time 4515.2 Seconds  
(without alignments)  
10582.571 Million cell updates/sec

Title: US-09-729-264-3  
Perfect score: 1168  
Sequence: 1 agtgatcaggggagcaggag.....gtatacaactgtagtatag 1168

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

SUMMARIES

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1168	100.0	1168	6	AX380398	Sequence
2	1122.4	96.1	1175	6	AX380396	Sequence
3	1119.6	95.9	2051	9	AK092516	Homo sapi
4	1119	95.8	1240	6	AX380400	Sequence
5	1102.8	94.4	1139	6	AX380402	Sequence
6	378.4	32.4	1195	6	AX380404	Sequence
7	376.2	32.2	1155	10	AF537215	Mus muscu
8	326.6	28.0	142742	9	AF121782	Homo sapi
9	326.6	28.0	340000	9	HS21C080	Homo sapi
10	182.4	15.6	149964	2	AC120145	Mus muscu
11	182.4	15.6	181510	2	AC120346	Mus muscu
12	182.4	15.6	196900	2	AC020851	Mus muscu
13	181.8	15.6	754	6	AX380408	Sequence
14	181.8	15.6	895	6	AX380406	Sequence
15	181.8	15.6	1556	10	BC004806	Mus muscu
16	152	13.0	47467	2	AC135456	Rattus no
17	152	13.0	140868	2	AC142238	Rattus no
18	151.4	13.0	181510	2	AC120346	Mus muscu
19	100	8.6	199665	9	AF064857	Homo sapi
20	57.4	4.9	196900	2	AC020851	Mus muscu
21	56.6	4.8	427	9	HSU63332	Human super
22	55.6	4.8	250956	2	AC118358	Rattus no
23	54	4.7	1061	9	HSCAGCTG	Homo sapi
24	54.8	4.7	149964	2	AC120145	Mus muscu
25	53	4.5	304785	2	AC118366	Rattus no
26	52.8	4.5	110000	3	AE003536	Continuation (3 of
27	52.8	4.5	138685	2	AC017582	Continuation (3 of
28	52.8	4.5	166249	3	AC093502	Drosophil
29	52.8	4.5	185404	3	AC104703	Drosophil
30	52	4.5	280780	2	AC125940	Rattus no
31	51.8	4.4	397	6	BD010479	Diagnosti
32	51	4.4	136171	2	AC141946	Rattus no
33	51	4.4	192820	2	AC102163	Mus muscu
34	50.6	4.3	147185	9	AC020589	Homo sapi
35	50.6	4.3	192292	2	AC129068	Rattus no
36	50.6	4.3	320082	2	AC113925	Rattus no
37	50	4.3	110000	2	AC108346	Continuation (2 of
38	50	4.3	226725	2	AC113770	Rattus no
39	50	4.3	237179	2	AC128466	Rattus no
40	50	4.3	244322	2	AC123135	Rattus no
41	49.8	4.3	299	9	HSAL31896	Homo sapi
42	49.8	4.3	209706	2	AC119611	Rattus no
43	49.8	4.3	227341	2	AC127771	Rattus no
44	49.8	4.3	231008	2	AC129166	Rattus no
45	49.8	4.3	286524	2	AC105802	Rattus no

ALIGNMENTS

RESULT 1  
AX380398  
LOCUS AX380398 1168 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 3 from Patent WO0200710.  
ACCESSION AX380398  
VERSION AX380398.1 GI:19575328  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS Welcher A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.  
TITLE B7-like molecules and uses thereof  
JOURNAL Patent: WO 0200710-A 3 03-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

```

FEATURES             Angen, Inc. (US)
source               Location/Qualifiers
                    1. .1168
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                    8. .1168
                      /note="unnamed protein product"
CDS
/codon_start=1
/protein_id="CAD28217.1"
/db_xref="GI:19575329"
/translation="MVAGANENRDPGSGNEVIEGPNARVLKGSQARFNCTVSG
WKLMALSDMVLSVRPMEPIITNDRFTSQRDQGNFTSEMIHNVPDSGNIR
SLONGSLHSAVLTQVWVMELEIPSNLVVAENECEVTCLPSHWTLPDLSWELGL
VHSSYFVPEPSDLQSAVLSIALTPQSGNLTCTVATKSLKARKSATVNLTIRCPQ
DTGGGNIPEVLSLPSLFTWKGVLGLAGTLMLLTPCTLTIRCCCRRCGCG
NCCRCFCCKRRKGFRIQFKKSEKTNKETETESNGNSGYNSDKQKTTDASLP
PKSCSSDPEQRNSGPPHQRADQPPRPASHPQASFNLASPEKVSNTTV"
BASE COUNT          299 a 302 c 288 g 279 t
ORIGIN
Query Match          100.0%; Score 1168; DB 6; Length 1168;
Best Local Similarity 100.0%; Pred. No. 3,2e-268; Indels 0; Gaps 0;
Matches 1168; Conservative 0; Mismatches 0;

Qy 1 AGTGATCATGTGCGCAGGAGCCATGAAATAGAGACCCACCGGTTCTGGGTCTGGTAA 60
Db 1 AGTGATCATGTGCGCAGGAGCCATGAAATAGAGACCCACCGGTTCTGGGTCTGGTAA 60
Qy 61 TGAAGTCATAGAGAGGCCCCCAAAATCAAGAGTCCTGAAGGCTCCAGGCTCGCTTCAA 120
Db 61 TGAAGTCATAGAGAGGCCCCCAAAATCAAGAGTCCTGAAGGCTCCAGGCTCGCTTCAA 120
Qy 121 CTGCACCGTCTCCAGGCGTGGAGGTCATCATGTGGGCTCTCAGTGACATGGTGGTCT 180
Db 121 CTGCACCGTCTCCAGGCGTGGAGGTCATCATGTGGGCTCTCAGTGACATGGTGGTCT 180
Qy 181 AAGGCTCAGGCGCATGAGCGCATCATCAGCATGACCGCTTCACTCTCAGAGTACGA 240
Db 181 AAGGCTCAGGCGCATGAGCGCATCATCAGCATGACCGCTTCACTCTCAGAGTACGA 240
Qy 241 CCAGGCGGGAACCTTCACTCGGAGATGATCATCCCAATGTGGAGCCAGTGATTCGG 300
Db 241 CCAGGCGGGAACCTTCACTCGGAGATGATCATCCCAATGTGGAGCCAGTGATTCGG 300
Qy 301 GAACATCAGATGACGCTCCAGACAGTCGCTGATGATGATGATGATGATGATGATGAT 360
Db 301 GAACATCAGATGACGCTCCAGACAGTCGCTGATGATGATGATGATGATGATGATGAT 360
Qy 361 AGTTATGGAGAGCTGTTATCCAGTGTATCTTGTAGTGGCTGAGAAATGAACCTTG 420
Db 361 AGTTATGGAGAGCTGTTATCCAGTGTATCTTGTAGTGGCTGAGAAATGAACCTTG 420
Qy 421 TGAAGTTACTTGTCTACCTCCACAGTGGAGCTGGCTCCGATATTTCTCGGAGCTGG 480
Db 421 TGAAGTTACTTGTCTACCTCCACAGTGGAGCTGGCTCCGATATTTCTCGGAGCTGG 480
Qy 481 TCTCCTGTCAGGCATTCAGCTATTATTTTGTTCGGAGCCAGCGACCTTCAAAGTGC 540
Db 481 TCTCCTGTCAGGCATTCAGCTATTATTTTGTTCGGAGCCAGCGACCTTCAAAGTGC 540
Qy 541 AGTGAGATCTGGCTCTGACCCACAGCATGGAGCTTTGACCTTGGCTGGCTACCTG 600
Db 541 AGTGAGATCTGGCTCTGACCCACAGCATGGAGCTTTGACCTTGGCTGGCTACCTG 600
Qy 601 GAAGAGCTGAAGCGCCGCAAGTCTGCACTGTAAATCTCACTGTGATTCGGTGTCCCA 660
Db 601 GAAGAGCTGAAGCGCCGCAAGTCTGCACTGTAAATCTCACTGTGATTCGGTGTCCCA 660
Qy 661 AGACACTGGAGCTGATTAATATCCAGGTGATTAATCAAGTTTACCGAGTTTAGTTT 720
Db 661 AGACACTGGAGCTGATTAATATCCAGGTGATTAATCAAGTTTACCGAGTTTAGTTT 720

```

```

Qy 721 TTCATTGCTCTACTTGGGCAAAAGTTGGACTTGGACTAGCAGCACCATGCTTCTGACGCC 780
Db 721 TTCATTGCTCTACTTGGGCAAAAGTTGGACTTGGACTAGCAGCACCATGCTTCTGACGCC 780
Qy 781 GAGCTGCTACTCTTACAAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 GAGCTGCTACTCTTACAAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 841 CTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 CTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 GAAATCTGAAAAGAGCAACACAAAGAACTCAGACAGAAAGTGAATAAGAAATCTC 960
Db 901 GAAATCTGAAAAGAGCAACACAAAGAACTCAGACAGAAAGTGAATAAGAAATCTC 960
Qy 961 CGGCTCAATTCAGATGAAACAAAGACCAACAGACACCGCTTCTCTCCCTCCCAATCT 1020
Db 961 CGGCTCAATTCAGATGAAACAAAGACCAACAGACACCGCTTCTCTCCCTCCCAATCT 1020
Qy 1021 TGAATCCAGTCTCTGAAACAAAGAACTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 TGAATCCAGTCTCTGAAACAAAGAACTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 TCAACGCTCCACCGAGCCAGCAAGTCCATCCACAGGCTTCTTTAATCTGGCCAGTCTG 1140
Db 1081 TCAACGCTCCACCGAGCCAGCAAGTCCATCCACAGGCTTCTTTAATCTGGCCAGTCTG 1140
Qy 1141 GAAGGTCAGTAAATACAACTGTAGTATAG 1168
Db 1141 GAAGGTCAGTAAATACAACTGTAGTATAG 1168

RESULT 2
AX380396          1175 bp      DNA      linear      PAT 18-MAR-2002
LOCUS             Sequence 1 from Patent WO0200710.
DEFINITION        AX380396
ACCESSION         AX380396.1 GI:19575326
VERSION           AX380396.1
KEYWORDS           Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE          1
AUTHORS            Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE              B7-like molecules and uses thereof
JOURNAL            Patent: WO 0200710-A 1 03-JAN-2002;
                  Angen, Inc. (US)
FEATURES           Location/Qualifiers
source            1. .1175
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
CDS               27. .1175
                  /note="unnamed protein product"
                  /codon_start=1
                  /protein_id="CAD28216.1"
                  /db_xref="GI:19575327"
                  /translation="MGLVFLHSGSGNEVIEGPNATVLKGSQARFNCTVSGWKLI
                  SWALSDMVLSVRPMEPIITNDRFTSQRDQGNFTSEMIHNVPDSGNIRCSLQN
                  SYFVPEPSDLQSAVLSIALTPQSGNLTCTVATKSLKARKSATVNLTIRCPQDTGG
                  GNIPGVLSLPSLFTWKGVLGLAGTLMLLTPCTLTIRCCCRRCGCGCC
                  RCCFCCKRRKGFRIQFKKSEKTNKETETESNGNSGYNSDKQKTTDASLPKSC
                  ESSDPEQRNSGPPHQRADQPPRPASHPQASFNLASPEKVSNTTV"
BASE COUNT        295 a 307 c 286 g 287 t
ORIGIN

Query Match       96.1%; Score 1122.4; DB 6; Length 1175;
Best Local Similarity 99.5%; Pred. No. 2.5e-257;
Matches 1126; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 37 CCCACCGGTTCTGGGTTCTGGTAATGAAGTCATAGAGGCCGCCAAAATGCAAGAGTCT 96  
 Db 44 CCTCCACGGTTCTGGGTTCTGGTAATGAAGTCATAGAGGCCGCCAAAATGCAAGTCT 103  
 QY 97 GAAGGGTCCAGGCTCGCTTAACTGACCGCTCTCCAGGGCTGGAAGCTCATCATCTG 156  
 Db 104 GAAGGGTCCAGGCTCGCTTAACTGACCGCTCTCCAGGGCTGGAAGCTCATCATCTG 163  
 QY 157 GGCTCTCAGTGACATGGTGTGTCTGAAGCGTCAGGCCCATGGAGCCCATCATCACCAATGA 216  
 Db 164 GGCTCTCAGTGACATGGTGTGTCTGAAGCGTCAGGCCCATGGAGCCCATCATCACCAATGA 223  
 QY 217 CCGTTCACCTCTCAGAGGTACACACAGGGCGGGAACCTTCACTCGGAGATGATCATCA 276  
 Db 224 CCGTTCACCTCTCAGAGGTACACACAGGGCGGGAACCTTCACTCGGAGATGATCATCA 283  
 QY 277 CAATGTGGAGCCCATGATTCGGGGAAACATCAGATCAGCGCTCCAGAACAGTCGCTGCA 336  
 Db 284 CAATGTGGAGCCCATGATTCGGGGAAACATCAGATCAGCGCTCCAGAACAGTCGCTGCA 343  
 QY 337 TGGATCTGCTTACCTTACCGTCCAGGTTATGGAGAGCTGTTTCATCCAGGTTTAATCT 396  
 Db 344 TGGATCTGCTTACCTTACCGTCCAGGTTATGGAGAGCTGTTTCATCCAGGTTTAATCT 403  
 QY 397 TGTAGTCGCTGAGAACTTGTGAAGTTACTTGTCTACCTCACACTGACCTGGCT 456  
 Db 404 TGTAGTCGCTGAGAACTTGTGAAGTTACTTGTCTACCTCACACTGACCTGGCT 463  
 QY 457 CCGGATATTCCTCGGAGCTCGGTCTCCCTGGTCAGCATTCAAGCTATTAATTTGTTC 516  
 Db 464 CCGGATATTCCTCGGAGCTCGGTCTCCCTGGTCAGCATTCAAGCTATTAATTTGTTC 523  
 QY 517 GGAGCCAGGACCTTCAAAGTGCAGTGAGCATCTCGCTCTGACCCCAAGCAATGG 576  
 Db 524 GGAGCCAGGACCTTCAAAGTGCAGTGAGCATCTCGCTCTGACCCCAAGCAATGG 583  
 QY 577 GACTTTCAGTTCGCTGCTTACCTGGAAGAGCTGAGGCGCGCACTGCGCACTGTAA 636  
 Db 584 GACTTTCAGTTCGCTGCTTACCTGGAAGAGCTGAGGCGCGCACTGCGCACTGTAA 643  
 QY 637 TCTCACTGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTAATATTCAGGTGAT 696  
 Db 644 TCTCACTGTGATTCGGTGTCCCAAGACACTGGAGGTGGTATTAATATTCAGGTGAT 703  
 QY 697 ATCAGTTTACGAGTTTGGTATTTTCAATGCTTCTGCGGCAAGTGGCTGGACT 756  
 Db 704 ATCAGTTTACGAGTTTGGTATTTTCAATGCTTCTGCGGCAAGTGGCTGGACT 763  
 QY 757 AGCAGGCACCATGTTCTGACGCGAGCTGCTTCTTACAAATACGCTGCTGTGCGG 816  
 Db 764 AGCAGGCACCATGTTCTGACGCGAGCTGCTTCTTACAAATACGCTGCTGTGCGG 823  
 QY 817 CCGTGTGTTGTTGGCTGCACTCTGCTGCGTGTGTTTCTGCTGTAGAGAAAG 876  
 Db 824 CCGTGTGTTGTTGGCTGCACTCTGCTGCGTGTGTTTCTGCTGTAGAGAAAG 883  
 QY 877 AGGATTCGTATTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAAACAAAGAACTGA 936  
 Db 884 AGGATTCGTATTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAAACAAAGAACTGA 943  
 QY 937 GACGAAAGTGAATGAAACTCCGCTCAATTCAGATGAAACAAAGAACACACACAC 996  
 Db 944 GACGAAAGTGAATGAAACTCCGCTCAATTCAGATGAAACAAAGAACACACACAC 1003  
 QY 997 CGCTTCTCTCCCTCCCAATCTGTGTAATCCAGTGATCTCGAACAAAGAAACAGTGTG 1056  
 Db 1004 CGCTTCTCTCCCTCCCAATCTGTGTAATCCAGTGATCTCGAACAAAGAAACAGTGTG 1063  
 QY 1057 TGGCCCTCTCAGAGCGGCTGATCAGCTCCACCGCCAGCTCATCCAGGC 1116  
 Db 1064 TGGCCCTCTCAGAGCGGCTGATCAGCTCCACCGCCAGCTCATCCAGGC 1123  
 QY 1117 TTCTTTTAACTCTGGCAGTCTGAGAGGTCAGTAATACAACTGTAGTATAG 1168

Db 1124 TTCTTTTAACTCTGGCAGTCTGAGAGGTCAGTAATCAACTGTAGTATAG 1175  
 RESULT 3  
 AK092516 2051 bp mRNA linear PRI 15-JUL-2002  
 LOCUS Homo sapiens cDNA FLJ35197 fis, clone PLACE6017788, highly similar  
 to IGSF5.  
 AK092516  
 DEFINITION AK092516.1 GI:21751130  
 ACCESSION AK092516  
 VERSION oligo capping; fis (full insert sequence).  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,  
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
 Kawai-Hio, Y., Saico, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
 Wagatsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,  
 Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.  
 and Isogai, T.  
 and Isogai, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2051)  
 REFERENCE Isogai, T. and Yamamoto, J.  
 Direct Submission  
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
 HRI, and Biotechnology Center, National Institute of Technology and  
 Evaluation; clone selection for full insert sequencing: HRI and  
 RAB; annotation: HRI and RAB.  
 FEATURES  
 source  
 1. 2051  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACE6017788"  
 /tissue\_type="placenta"  
 /clone\_lib="PLACE6"  
 /notes="cloning vector: pME18SFL3"  
 BASE COUNT 580 a 463 c 474 g 534 t  
 ORIGIN  
 Query Match 95.9%; Score 1119.6; DB 9; Length 2051;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-256; Indels 0; Gaps 0;  
 Matches 1122; Conservative 0; Mismatch 4;  
 QY 43 CGGTTCTGGGTTCTGGTAATGAAGTCATAGAGGCCGCCAAAATGCAAGAGTCTGAGGG 102  
 Db 202 CGGTTCTGGGTTCTGGTAATGAAGTCATAGAGGCCGCCAAAATGCAAGAGTCTGAGGG 261  
 QY 103 CTCCCAGGCTCGCTTCAACTGCACTGCAAGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCT 162  
 Db 262 CTCCCAGGCTCGCTTCAACTGCACTGCAAGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCT 321  
 QY 163 CAGTGACATGGTGTCTGAAGCGTCAAGCGCCATGGAGCCCATCATCCCAATGACCGTT 222  
 Db 322 CAGTGACATGGTGTCTGAAGCGTCAAGCGCCATGGAGCCCATCATCCCAATGACCGTT 381  
 QY 223 CACCTCTCAGAGTACGACCGCGGGAACCTTCACTCGGAGTGTATCATCAAGT 282  
 Db 382 CACCTCTCAGAGTACGACCGCGGGAACCTTCACTCGGAGTGTATCATCAAGT 441  
 QY 283 GGAGCCAGTGTATTGGGGGAACATCAGATCGAGCTCCAGAACAGTGCCTGCTGATGATC 342

Db 442 GGAGCCAGTATTCGGGAAACATCAGATCCAGCTCCAGAACAGTCGCTGCATGATC 501  
 Qy 343 TGCTTACCTTACCGTCCAAAGTTATGAGAGAGCTGTTTCATTCAGTGTATTCCTTGTAGT 402  
 Db 502 TGCTTACCTTACCGTCCAAAGTTATGAGAGAGCTGTTTCATTCAGTGTATTCCTTGTAGT 561  
 Qy 403 CGCTGAGAATGAACCTTGTGAAGTTACTTGTCTTACCTCACAAGCTGAGCTGCTCCCGA 462  
 Db 562 CGCTGAGAATGAACCTTGTGAAGTTACTTGTCTTACCTCACAAGCTGAGCTGCTCCCGA 621  
 Qy 463 TATTTCTGGAGCTCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522  
 Db 622 TATTTCTGGAGCTCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681  
 Qy 523 CAGCGACCTTCAAAGTGCAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582  
 Db 682 CAGCGACCTTCAAAGTGCAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741  
 Qy 583 GACTTGGTGGTACCTGGAAGAGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGG 642  
 Db 742 GACTTGGTGGTACCTGGAAGAGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGG 801  
 Qy 643 TGTGATTCGGTGTCCCAAGACATCGAGTGTGTATTAATATTCAGGTGTATTAATATTCAGG 702  
 Db 802 TGTGATTCGGTGTCCCAAGACATCGAGTGTGTATTAATATTCAGGTGTATTAATATTCAGG 861  
 Qy 703 TTTTACCGAGTTAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762  
 Db 862 TTTTACCGAGTTAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921  
 Qy 763 CACCATGCTTCTGAGCGGCGAGTGTACTCTTAACAATACGCTGCTGCTGCTGCTGCTGCTGCT 822  
 Db 922 CACCATGCTTCTGAGCGGCGAGTGTACTCTTAACAATACGCTGCTGCTGCTGCTGCTGCTGCT 981  
 Qy 823 TTGTTGGGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882  
 Db 982 TTGTTGGGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041  
 Qy 883 TCGTATTCATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAAACTGAGACAGA 942  
 Db 1042 TCGTATTCATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAAACTGAGACAGA 1101  
 Qy 943 AAGTGAATGAAATCTCCGGCTACATTCAGATGAACAAAGACACACAGACACCGCTTC 1002  
 Db 1102 AAGTGAATGAAATCTCCGGCTACATTCAGATGAACAAAGACACACAGACACCGCTTC 1161  
 Qy 1003 TCTCCCTCCCAATCTCTGCAATCCAGTGAATCCAGTGAATCCAGTGAATCCAGTGAATCCAGT 1062  
 Db 1162 TCTCCCTCCCAATCTCTGCAATCCAGTGAATCCAGTGAATCCAGTGAATCCAGTGAATCCAGT 1221  
 Qy 1063 TCCCTCAGCAGGGGCTGATCAACGCTCCAGCAGGCGAGCAAGTCAATCCAGTGAATCCAGTGAAT 1122  
 Db 1222 TCCCTCAGCAGGGGCTGATCAACGCTCCAGCAGGCGAGCAAGTCAATCCAGTGAATCCAGTGAAT 1181  
 Qy 1123 TAATCTGCCAGTCTCAGAGGCTCAGTAATACATGATGATATAG 1168  
 Db 1282 TAATCTGCCAGTCTCAGAGGCTCAGTAATACATGATGATATAG 1327

RESULT 4  
 AX380400  
 LOCUS AX380400 1240 bp DNA linear PAT 18-MAR-2002  
 DEFINITION Sequence 5 from Patent WO0200710.  
 ACCESSION AX380400  
 VERSION AX380400.1 GI:19575330  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Welcher, A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.

B7-like molecules and uses thereof  
 Patent: WO 0200710-A 5 03-JAN-2002;  
 Amgen, Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..1240  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 80..1240  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD28218.1"  
 /db\_xref="GI:19575331"  
 /translation="MERHLLTTPVAVGSGSNEVLEGPNATVLKGSQARNCTVSGQ  
 WKUIMWALSDMVLVSVRMEPIITNDRFTSQIDOGNFTSEMILHNVPDSGNIRL  
 SLONSLHGSAYLTVOVMGELFIPSNVLVAENECEFTCLPSHWRULPDLISWEGLL  
 VHSYFYFVPEPSDLSAVSIALTLTQSNGLTLCVATKSLKARSAATNLTVCRCQ  
 DTGGSLNIPFVLSPLGSLGTWGLAGLMLTPTCTLTIRCCRRRCQCG  
 NCCRCFCFRRRGFRIOPOKKSEKTKETETESGNGNSGYNSDKQKTTETASLP  
 PKSCSSDPEQRNSSCGPHQRADORPPRPASHPOASFNLASPEKVSNTTVV"  
 BASE COUNT 318 a 319 c 305 g 298 t  
 ORIGIN

Query Match 95.8%; Score 1119; DB 6; Length 1240;  
 Best Local Similarity 99.1%; Pred. No. 1.6e-256;  
 Matches 1125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Qy 34 AGACCCACCGGTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 93  
 Db 106 AGAGCTGTAGGTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 165  
 Qy 94 CTTGAAGGGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 153  
 Db 166 CTTGAAGGGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225  
 Qy 154 GTGGGCTCTCAGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213  
 Db 226 GTGGGCTCTCAGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285  
 Qy 214 TGACCGGCTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 273  
 Db 286 TGACCGGCTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345  
 Qy 274 CCAATATGTCGAGCCAGTGTTCGGGGACACATCAGATGACGCTCCAGAACAGTCCGCT 333  
 Db 346 CCAATATGTCGAGCCAGTGTTCGGGGACACATCAGATGACGCTCCAGAACAGTCCGCT 405  
 Qy 334 GCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGTAA 393  
 Db 406 GCATGGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGTAA 465  
 Qy 394 TCTTGTAGTCGCTGAGATGACCTTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453  
 Db 466 TCTTGTAGTCGCTGAGATGACCTTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525  
 Qy 454 GCTCCCGGATATTTCTGGAGCTCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513  
 Db 526 GCTCCCGGATATTTCTGGAGCTCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585  
 Qy 514 TCCGGAGCCAGCGACCTTCAAGTGCAGTGCATCTCTGGCTCTGACCCACAGAGCAA 573  
 Db 586 TCCGGAGCCAGCGACCTTCAAGTGCAGTGCATCTCTGGCTCTGACCCACAGAGCAA 645  
 Qy 574 TGGGACTTTTACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633  
 Db 646 TGGGACTTTTACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705  
 Qy 634 AATATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693  
 Db 706 AATATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765  
 Qy 694 AATATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753

```

Db      766  ATTATCAAGTTTACCGAGTTTAAAGTTTTCATTTGTCCTACTTGGGGCAAAAGTTGGACTTGG 825
Qy      754  ACTAGCAGGACCATGCTCTGACGCGCAGCTGTACTCTTACAATACGCTGCTGCTGCTG 813
Db      826  ACTAGCAGGACCATGCTCTGACGCGCAGCTGTACTCTTACAATACGCTGCTGCTGCTG 885
Qy      814  CCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 873
Db      886  CCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
Qy      874  AAGAGGATTTGCTATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAAAC 933
Db      946  AAGAGGATTTGCTATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAAAC 1005
Qy      934  TGAGACAGAGTGGAAATGAAATCCGCTACAAATTCAGATGAAACAAAGACACACAGA 993
Db      1006  TGAGACAGAGTGGAAATGAAATCCGCTACAAATTCAGATGAAACAAAGACACACAGA 1065
Qy      994  CACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1053
Db      1066  AACCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1125
Qy      1054  CTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1113
Db      1126  CTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1185
Qy      1114  GGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1168
Db      1186  GGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1240

RESULT 5
AX380402
LOCUS      AX380402                1139 bp      DNA      linear      PAT 18-MAR-2002
DEFINITION Sequence 7 from Patent WO0200710.
ACCESSION  AX380402
VERSION     AX380402.1  GI:19575332
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Weiher, A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
  TITLE     B7-like molecules and uses thereof
  JOURNAL   Patent: WO 0200710-A 7 03-JAN-2002;
            Amgen, Inc. (US)
FEATURES
  source    Location/Qualifiers
            1..1139
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
            1..1134
              /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD28219.1"
            /db_xref="GI:19575333"
            /translation="MVGAGMENDPPGSGSNEVIEGPQNRVLKGSQRFNCTVSQG
            WKLTWALSQMVLSVRMPEPIITNDRFTSQRYDQGNLTSEMIHNVEPSDGNIRC
            SLQNSRLHSAYLTVQVMBELFIPSVNLVVAENEPCVETCLPSHWRLPDIISWELGLL
            VSHSYVFPVPSDLQSAVSLALTPQSNGLTCTVATWKSARKSATVNLTVIRCPQ
            DTGGGINIPGVLSPLSGFLSIPTWKVGVLGAGTMLTPTCTLTIRCCCRRCRCG
            NCCRCFCPCRRKRNKLRKOTKLRKVRKEMTPATIQMKRPQTFLLSLPNPNPV
            ILNKETVAVALLTSLINVHPGOQVIRHLLLIWPLVLRSSVQL"
BASE COUNT 290 a 300 c 283 g 266 t
ORIGIN
Query Match 94.4%; Score 1102.8; DB 6; Length 1139;
Best Local Similarity 97.9%; Pred. No. 1.2e-252;
Matches 1136; Conservative 0; Mismatches 2; Indels 22; Gaps 1;
Qy      8  ATGTGGGAGGCGCATGAAATAGAGACCCACCGCTTCTGGGCTCGTATGAATC 67

```

```

Db      1  ATGCTGGCAGGAGCCATGGAATAATAGAGACCAACCCGGTCTGGGTCTGGTAAATGAAGTC 60
Qy      68  ATAGAAGGCCCCCAAAATGCAAGAGTCTTGAAGGCTCCCAAGGCTGCTTCAACTGCAACC 127
Db      61  ATAGAAGGCCCCCAAAATGCAAGAGTCTTGAAGGCTCCCAAGGCTGCTTCAACTGCAACC 120
Qy      128  GTCTCCACAGGCTGGAGGCTCATGTTGGGCTCTCAGTGATGATGGTGGTCTTAAGCGTC 187
Db      121  GTCTCCACAGGCTGGAGGCTCATGTTGGGCTCTCAGTGATGATGGTGGTCTTAAGCGTC 180
Qy      188  AGGCCCATGAGAGCCCATCATCAATGACCCCTTCACTCTCAGAGTACGACAGGCGC 247
Db      181  AGGCCCATGAGAGCCCATCATCAATGACCCCTTCACTCTCAGAGTACGACAGGCGC 240
Qy      248  GGAACCTTCACCTCGAGATGATCATCAATGAGGCCCGAGTATTCGGGGAAATC 307
Db      241  GGAACCTTCACCTCGAGATGATCATCAATGAGGCCCGAGTATTCGGGGAAATC 300
Qy      308  AGATGAGGCTCCAGAACAGTCCCTGCAATGATCTGCTTACCTTACCGTCCAGTTATG 367
Db      301  AGATGAGGCTCCAGAACAGTCCCTGCAATGATCTGCTTACCTTACCGTCCAGTTATG 360
Qy      368  GAGAGCTGTTCAATCCCAAGTGTAAATCTTGTAGTCTGAGTGAAGTGAACCTTGAAGTT 427
Db      361  GAGAGCTGTTCAATCCCAAGTGTAAATCTTGTAGTCTGAGTGAAGTGAACCTTGAAGTT 420
Qy      428  ACTTGTCTACCTCACAACCTGGCTCCCGGATATTTCTGGAGCTGGTCTCTCTG 487
Db      421  ACTTGTCTACCTCACAACCTGGCTCCCGGATATTTCTGGAGCTGGTCTCTCTG 480
Qy      488  GTGAGCATTCAAGCTTATTTTGTTCGGAGCCCGAGGACCTTCAAGTCACTGAGC 547
Db      481  GTGAGCATTCAAGCTTATTTTGTTCGGAGCCCGAGGACCTTCAAGTCACTGAGC 540
Qy      548  ATCTGCTCTGACCCACAGAGCAATGGACTTTGACTTGGTGGCTACCTGGAAGAGC 607
Db      541  ATCTGCTCTGACCCACAGAGCAATGGACTTTGACTTGGTGGCTACCTGGAAGAGC 600
Qy      608  CTGAGGCCCGCAAGTCTGCAACTGTAAATCTCAGTGTGATTCGGTGTCCCAAGACACT 667
Db      601  CTGAGGCCCGCAAGTCTGCAACTGTAAATCTCAGTGTGATTCGGTGTCCCAAGACACT 660
Qy      668  GGAGGTGGTATTAATATTCAGGTGTATTAATCAAGTTTACCAAGTTTGGTTCATTG 727
Db      661  GGAGGTGGTATTAATATTCAGGTGTATTAATCAAGTTTACCAAGTTTGGTTCATTG 720
Qy      728  CCTACTTGGGCAAGTCTGCACTTGGACTGAGCAGGACCATCTTCTGAGCGCAGCTGT 787
Db      721  CCTACTTGGGCAAGTCTGCACTTGGACTGAGCAGGACCATCTTCTGAGCGCAGCTGT 780
Qy      788  ACTCTTACATACGCTGCTGCTGCTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
Db      781  ACTCTTACATACGCTGCTGCTGCTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      848  CGTTGTTGTTTCTGCTGTAGAGAAAAGAGGATTTGTTATTAATTTCAAAAGAAATCT 907
Db      841  CGTTGTTGTTTCTGCTGTAGAGAAAAGAGGATTTGTTATTAATTTCAAAAGAAATCT 878
Qy      908  GAAAAGAGAGAGCAAAACAAAGAAATCTGAGACAGAAAGTGAAGTGAAGTGAAGTGA 967
Db      879  GAAAAGAGAGAGCAAAACAAAGAAATCTGAGACAGAAAGTGAAGTGAAGTGAAGTGA 938
Qy      968  AATTCAGATCAACAAAGAGACCAAGACCGCTTCTCTCCCTCCCAAAATCTGTAATCC 1027
Db      939  AATTCAGATCAACAAAGAGACCAAGACCGCTTCTCTCCCTCCCAAAATCTGTAATCC 998
Qy      1028  AGTGATCTCTCAACAAAGAAACAGTAGCTGTGGCCCTCTCTCACCAGCGGCTGATCAAGT 1087
Db      999  AGTGATCTCTCAACAAAGAAACAGTAGCTGTGGCCCTCTCTCACCAGCGGCTGATCAAGT 1058
Qy      1088  CCACCCAGGCGGAGAGTCAATCCACAGGCTTCTTTTAATCTCGGAGCTCTGAGAGGTC 1147
Db      1059  CCACCCAGGCGGAGAGTCAATCCACAGGCTTCTTTTAATCTCGGAGCTCTGAGAGGTC 1118

```

QY 1148 ACTAATACAACTGTAGTATA 1167  
 Db 1119 AGTAATACAACTGTAGTATA 1138  
  
 RESULT 6  
 LOCUS AX380404 1195 bp DNA linear PAT 18-MAR-2002  
 DEFINITION Sequence 9 from Patent WO0200710.  
 ACCESSION AX380404  
 VERSION AX380404.1 GI:19575334  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 REFERENCE  
 1  
 AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.  
 TITLE B7-like molecules and uses thereof  
 JOURNAL Patent: WO 0200710-A 9 03-JAN-2002;  
 Angen, Inc. (US)  
 FEATURES  
 Location/Qualifiers  
 1..1195  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 53..1165  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD28220.1"  
 /db\_xref="GI:19575335"  
 /translation="MEGSRDVLAVLVLAQLTAGSSVQIIEGPQNVTVLKSEAHF  
 NCTVTHGMLMTLQNVLSLTQGFIIITNNRTYASVNSTSPISLIHHDVPS  
 DSGVCSQLNSHGFSALSVQVMGTNLIPSNLIVTEGFCNVTCVAGWTSLPDI  
 SWLEVEVSHSSVNSLEPNRVLVLDLPLGNLTCTVLAELKDLQASKSLTNL  
 TVVQPPDSIGEGPALPTWAILLAFAFILLIILVLIIFCCCAASREKESTY  
 QNEIRKASNRNKADPTWIKLKENYGVSSDEAKAQTASLPKPSAEVSLPEKRS  
 SLFYSINKHQGPATPRVPSFDIASPQVRNVLV"  
 BASE COUNT 313 a 304 c 291 g 287 t  
 ORIGIN  
 Query Match 32.4%; Score 378.4; DB 6; Length 1195;  
 Best Local Similarity 62.3%; Pred. No. 1.5e-79;  
 Matches 701; Conservative 0; Mismatches 356; Indels 69; Gaps 4;  
  
 QY 42 CCGGTTCTGGTCTGTAATGAGTCTATGAGGCGCCCAAAATCGAGAGTCCTGGAAG 101  
 Db 108 CAGCTTCCGGATCAGTTATCAGATCATGAGAGGCTCTCAGATGTAACAGTCTTAAAG 167  
  
 QY 102 GCTCCAGGCTCGCTTCAACTGACACCTGTCCAGGGCTGGAAGCTCATCATGTGGGCTC 161  
 Db 168 ACTCAGAGGCTCACTTCAACTGACACCTGTGACTCAGCGCTGGAAGCTTCTCATGTGAGCTC 227  
  
 QY 162 TCAGTGACATGGTGTCTGAAGCTGAGGCGCCATGAGCCATCATCAACATGACCGCT 221  
 Db 228 TTAACCAATGGTGTGTGAGTCTCACACCAAGAGCCATCATCAACACACCGCT 287  
  
 QY 222 TCACCTCTCAGAGTACGACCGGGGAACTTCACCTCGGAGATGATTCACCAATG 281  
 Db 288 TCACCTATGCCAGTTACACACAGCACTGACAGCTTCATCTCGGAGTTGATCATCCATGATG 347  
  
 QY 292 TGGAGCCAGGATTCGGGNAATCATGATGAGCTCCAGACAGCTCCGCTGCATGGAT 341  
 Db 348 TGCAGCCAGGATTCGGGATCCGATCGCAATGAGCTTCGAGACAGCCATGAGTTGAT 407  
  
 QY 342 CTGCTTACCTTACCGTCCAGTTATGGAGAGCTGTTCATTCCTCCAGTGTAAATCTGTAG 401  
 Db 408 CTGCTTACCTTACCGTCCAGTTCATGGGAGCCCTGAAATTCCTCAACACACCTTATAG 467  
  
 QY 402 TCGTGTGAATGAACCTTGTGAATCTTACTGTCTACCTCAGCTGAGCTGGCTCCCGG 461  
 Db 468 TCACTGAGGTTAAACCTTGTAAATGTGACTTGCTATGCGGTGGCTGAGCTCACTCCCGG 527

QY 462 ATATTTCTGGAGCTCGGTCTCTCTGTGAGCCATTCAAGCTATTATTTTGTTCGGAGC 521  
 Db 528 ATATTTCTGGAGCTTGAGGTTCCCGTAAGCCATTGAGTTACAATTTCTTCTGGAGC 587  
  
 QY 522 CAGCGACCTTCAAGTGCAGTGCCTCTGACCCACAGAGCAATGGGACTT 581  
 Db 588 CGGGCAACTTATGAGGCTTGTGAGCTTCAACACCTGAGGCAAGCGGACCT 647  
  
 QY 582 TGACTTGGTGGCTACCTGGAAGAGCTGGAAGCCCGCAAGTCTGCAACTGTAAATCTCA 641  
 Db 648 TGACTTGGTGGCAGAGCTGGAAGAGCTTGCAGGCGCAAGAGTCTTAACTGTCAAGCTGA 707  
  
 QY 642 CTGTGATTGGTGTCCCAAGACACTGGAGTGTGTAATTAATATTCAGAGGTGTAATCA 701  
 Db 708 CTGTGATTGGTGTCCCAAGACACTGGAGTGTGTAATTAATATTCAGAGGTGTAATCA 744  
  
 QY 702 GTTTACCGAGTTTGGTGTCTTCACTTGGGCAAGTGGACTTGGACTGGACTAGCAG 761  
 Db 745 -----AGGCCAGCACTGCGGCACTGCGGCACTGCGGCACTGCGGCACTGCGGCT 791  
  
 QY 762 GCACCATGCTTCTGAGCGCGGAGCTGTAATCAATAGCTGCTGCTGCGCGCGCTC 821  
 Db 792 TTTCTTCTGCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 843  
  
 QY 822 GTTGTGTGGTGCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881  
 Db 844 -----TTGTGCTCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898  
  
 QY 882 TTCTGATTCAATTTCAAAAGAAATCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 941  
 Db 899 TCTGAAACATGAG 944  
  
 QY 942 AAGTGGAAATGAAATCTCGGCTGATCAATCAGATGAAAGAGAGAGAGAGAGAGAGAG 1001  
 Db 945 AAGTGGAAATGAAATCTCGGCTGATCAATCAGATGAAAGAGAGAGAGAGAGAGAGAG 1004  
  
 QY 1002 CTCTCTCTCCCAATCTGTGATCCAGTGTGATCTGAAAGAGAGAGAGAGAGAGAGAG 1061  
 Db 1005 CTCTCTCTCTCAATCTGTGATCCAGTGTGATCTGAAAGAGAGAGAGAGAGAGAGAG 1058  
  
 QY 1062 CTCTCTCAGCGGGCTGATCAAGCTCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1121  
 Db 1059 TTTCTTATCAGGAAGTCAATTAACATCAGCGCGGCTCCAGCAACTCATCCAGGGTTTCT 1118  
  
 QY 1122 TTAATCTGCGGCTGCTGAGAGAGTCAATTAACATCTGATGATATA 1167  
 Db 1119 TTGACATCGCCAGTCTCTCAGAGGTCAGAGATGTGACTTTAGTATA 1164  
  
 RESULT 7  
 AF537215 1155 bp mRNA linear ROD 03-JUN-2003  
 LOCUS Mus musculus cell adhesion molecule JCAM mRNA, complete cds.  
 DEFINITION AF537215  
 ACCESSION AF537215.1 GI:31339733  
 VERSION AF537215.1  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1155)  
 REFERENCE 1  
 AUTHORS Hirabayashi, S., Tajima, M., Yao, I., Nishimura, W., Mori, H. and Hata, Y.  
 TITLE JAM4, a Junctional Cell Adhesion Molecule Interacting with a Tight Junction Protein, MAGI-1  
 JOURNAL Mol. Cell. Biol. 23 (12), 4267-4282 (2003)  
 MEDLINE 22658521  
 PUBMED 12773569  
 REFERENCE 2 (bases 1 to 1155)  
 AUTHORS Hata, Y., Hirabayashi, S. and Tajima, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-AUG-2002) Medical Biochemistry, Tokyo Medical and

```

Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan
Location/Qualifiers
1. 1155
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Swiss Webster"
/db_xref="taxon:10090"
43_1155
/notes="interacts with MAGI-1; localized at tight junctions
in epithelial cells"
/codon_start=1
/product="cell adhesion molecule JCAM"
/protein_id="AAP49218.1"
/db_xref="GI:31339734"
NCTTHGKLMWTTLNQVLSLTTOGPIITNNRFTYASYNSTDFISELIHVDQPH
/translations="MEGSRDVLAVILVLAQLTAGSSYQIIIEGPNVTVLKDSEAHF
DSGVSQCLONSHGFGAFSLVQVMGTLANIPNNLIVTBGECNVCTCAVGMTSLPDI
SWELEVPVSHSVNSFLESGNFWLSVLDLPLNGTLTCAVELKDAQSKSLVNL
TVQPPDSIGREGPALPTWAILLAVASILLIILVLIIFCCCSARREKSTY
QNEIRKANRKNKADPETKLGKKNYGHSDKAQAQASLPPKSAEVSUPEKRS
SLPYELNKHQFQPATHPRVRSFDIAPQKVRNVLIV"
BASE COUNT 299 a 296 c 283 g 277 t
ORIGIN
Query Match 32.2%; Score 376.2; DB 10; Length 1155;
Best Local Similarity 62.1%; Pred. No. 5.1e-79;
Matches 700; Conservative 0; Mismatches 358; Indels 69; Gaps 4;
QY 42 CCGTTCCTGGTCTGGTAATGATCATAGAGCCGCCAAATGCAAGAGTCTCTGAAG 101
DB 98 CAGCTTCGGGATCCAGTTATCAGATCATAGAGGTCCTCAGATGTAACAGTCTTAAG 157
QY 102 GTCCTCAGGCTCGCTTCAACTGACCGCTCTCCAGGCTGGAAGCTCATCATGTGGGCTC 161
DB 158 ACTCAGAGCTCACTTCACTGACCGGTGACTCAGGCTGGAAGCTTCTCATGTGGACTC 217
QY 162 TCACTGATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 221
DB 218 TTAACCAAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 277
QY 222 TCACTCTCAGAGGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 281
DB 278 TCACTATGCGGATTTACACAGCAGTACAGCTTCACTCGGAGTTGATCATCATGATG 337
QY 282 TGAGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 341
DB 338 TGCAGCCAGTACCTCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 397
QY 342 CTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 401
DB 398 CTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 457
QY 402 TCGCTGAGATGAACCTTGTGAGTACTTGTCTTACCTTACCTTACCTTACCTTACCTTAC 461
DB 458 TCACTGAGGAGTGAACCTTGTGAGTACTTGTCTTACCTTACCTTACCTTACCTTACCT 517
QY 462 ATATTTCCTGGAGCTCGGTCTCTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 521
DB 518 ATATTTCCTGGAGCTCGGTCTCTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 577
QY 522 CCAGGAGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 581
DB 578 CGGGCAACTTTATGAGGCTTTGAGTGTCTTGGACCTTCAACCTTGGGCAACGGGACCT 637
QY 582 TGACTTGTGCTGCTACCTGGAGAGCCTGGAAGGCGGCGGAGCTCTGCAACTGTAAATCTCA 641
DB 638 TGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 697
QY 642 CTGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 701
DB 698 CTGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
QY 702 GTTTACCGAGTTTAGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 761

```

```

Db 735 -----AGGCCAGCACTGCGGCACTCGGGCCATCATCTGCTGCGAGTGGCCT 781
QY 762 GCACCATGCTTCTGAGCGCGAGTGTACTCTTACAATACGCTGCTGCTGCTGCGCGCTC 821
Db 782 TTTCTTCTGCTTGTGATCCGATCATGTTTGAATTAATTAATTAATTAATTAATTAATTA 833
QY 822 GTTGTGCTGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
Db 834 -----TTGTGCTCTCCAGGAGAGAAAGGAAGAAATCTACTTATCAAAATGAATGA 888
QY 882 TTGCTATTCAATTTCAAAAGAAATCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 941
Db 889 TCTGCAAAATGAGGACAA-----ACAAAGAGAGATCCGGAGAGAGAGAGAGAGAG 934
QY 942 AAAGTGGAAATGAAACTCCGCTCAATTCAGATGAAACAAAGAGAGAGAGAGAGAGAG 1001
Db 935 AAGGTGAAAGGAAACTACGGGTACAGTTCGATGAGGAGAGAGAGAGAGAGAGAGAGAG 994
QY 1002 CTCTCCCTCCCAATCTCTGATCCAGTGTCTGATCCAGTGTCTGATCCAGTGTCTGAT 1061
Db 995 CTCTCCCTCTCAATCTCTGATCCAGTGTCTGATCCAGTGTCTGATCCAGTGTCTGAT 1048
QY 1062 CTCTCACCAGCGGCTGATCAACCTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1121
Db 1049 TTTCTTATCAGGAATCAATAAATCAATCAGCGGTCCAGCACTCATCCAGGAGTTTCT 1108
QY 1122 TTAATCTGCGCCTCTGAGAGAGTTCAGTAAATCAACAACTGTATGATAG 1168
Db 1109 TTGATCTGCGCCTCTCTGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 1155

RESULT 8
AF121782 142742 bp DNA linear PRI 02-FEB-1999
LOCUS Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.
DEFINITION AF121782
ACCESSION AF121782
VERSION AF121782.1 GI:4210991
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142742)
AUTHORS Taudien,S., Dagand,E., Hildmann,T., Nordseik,G., Drescher,B.,
Schatevov,R., Weber,D., Schilling,M., Menzel,U., Yaspo,M.L., and
Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source 1. 142742
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 206A10"
3. 118
/evidence=not experimental
repeat_region /rpt_family="MIR"
complement(140..415)
evidence=not experimental
repeat_region /rpt_family="MLT2E"
complement(668..881)
evidence=not experimental
repeat_region /rpt_family="MER58A"
complement(1078..1162)
/notes="GRAIL score = 86.000%, comment = excellent shadow"
evidence=not experimental
1114..1403
evidence=not experimental
repeat_region /rpt_family="MLT1I"

```

repeat_region	1481..1603 /evidence=not_experimental /rpt_family="I2"	repeat_region	9966..10041 /note="homology = 100.00%, score = 38, counts = 2"
exon	1576..1710 /note="MZEF, score = 66.18"	repeat_region	/evidence=not_experimental /rpt_type=tandem
exon	1576..1710 /evidence=not_experimental	repeat_region	/rpt_unit=tctctgcacactctctgtgagaggtcggtttatcgc
repeat_region	1576..1710 /note="MZEF, score = 66.18"	repeat_region	10244..10664 /evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="MLTIE"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	10665..11028 /evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="MLTIF"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	11318..11417 /evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="LIMC/D"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	11503..11818 /evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(11525..11777)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/note="MZEF, score = 79.48"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(12346..12827)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(12346..12827)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="MLTIC"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	12902..13180 /evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="AluSg"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(13057..13161)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/note="MZEF, score = 76.18"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(13226..13364)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(13226..13364)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="LIMC2"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(13613..13908)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="AluSx"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(14317..14614)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="LTR37A"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(14962..15036)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/note="MZEF, score = 938"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(15255..15299)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(15469..15497)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/note="Xpound exon prediction, score = 62% (0%)"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(16045..16768)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="MER21A"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(16555..16655)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/note="GRAIL, score = 54.000%, comment = good"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(16823..16850)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(16851..17208)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="THE1A"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(17209..17632)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="MER34"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(17633..17969)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="AluSg1"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(18004..18334)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="LTR17"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(18335..18632)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/evidence=not_experimental
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/rpt_family="HERV17"
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	complement(18352..18420)
repeat_region	1576..1710 /evidence=not_experimental	repeat_region	/note="GRAIL, score = 51.000%, comment = good shadow"

```

/existence=not_experimental
complement(18397..18584)
/notes="GRAIL, score = 66.000%, comment = good"
/evidence=not_experimental
complement(18643..24647)
/evidence=not_experimental
/rpt_family="HERV17"
complement(19228..19361)
/notes="WZEF, score = 92.8%"
/evidence=not_experimental
20900..21135
/notes="WZEF, score = 51.4%"
/evidence=not_experimental
21150..21275
/notes="Xpound exon prediction, score = 74% (0%)"

Query Match      28.0%; Score 326.6; DB 9; Length 142742;
Best Local Similarity 97.4%; Pred. NO. 4.7e-67;
Matches 332; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 27 AATAAGAGACCCACCGGCTTCGGTCTGGTAATGAGTCATAGAGGCCGCCAAAATG 86
Db 44257 AATTTGGTACCTCCAGGCTTCGGTCTGGTAATGAGTCATAGAGGCCGCCAAAATG 44316

Qy 87 CAAGAGTCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTCGAAGC 146
Db 44317 CAAGAGTCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTCGAAGC 44376

Qy 147 TCATCATGTGGGCTTCAGTGACATGCTGTCTAGCGTCAGGCGCCATGGAGCCCATCA 206
Db 44377 TCATCATGTGGGCTTCAGTGACATGCTGTCTAGCGTCAGGCGCCATGGAGCCCATCA 44436

Qy 207 TCACCAATGACCGCTTCACTCTCAGAGGTACGACAGGCGGGGAACTTCACTCGGAGA 266
Db 44437 TCACCAATGACCGCTTCACTCTCAGAGGTACGACAGGCGGGGAACTTCACTCGGAGA 44496

Qy 267 TGATCATCCACATGTGGGCGGAGTTCGGGGACATCAGATCGAGCTCCAGGACA 326
Db 44497 TGATCATCCACATGTGGGCGGAGTTCGGGGACATCAGATCGAGCTCCAGGACA 44556

Qy 327 GTGCGCTGCATGATCTGCTTACCTTACCGTCCAAAGTTATG 367
Db 44557 GTGCGCTGCATGATCTGCTTACCTTACCGTCCAAAGTTATG 44597

RESULT 9
LOCUS HS21C080 340000 bp DNA linear PRI 24-MAY-2000
DEFINITION Homo sapiens chromosome 21 segment HS21C080.
ACCESSION AL163280 AP001735 BA000005
VERSION AL163280.2 GI:7717369
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 340000)
Park H.S., Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E.,
Okiki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
Polley A., Wenzel U., Delabar J., Kumpf K., Lehmann R.,
Patterson D., Reichwald K., Rump A., Schillabel M., Schudy A.,
Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K.,
Antonarakis S.E., Minoshima S., Shimizu N., Nordliek G.,
Hornischer K., Brandt P., Scharte M., Schoen O., Desario A.,
Reichelt J., Kauer G., Bloeker H., Ramser J., Beck A., Klages S.,
Hennig S., Rieselmann L., Dagand E., Wehmeyer S., Borzym K., and
Gardiner K., Nizetic D., Francis F., Lehrach H., Reinhardt R., and
Yaspo M.L.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research

```

COMMENT

```

Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagami-hara 228-8555, Japan,
* e-mail: sakaki@gs.c.riken.go.jp/
* URL: http://hgp.gsc.riken.go.jp/
and
* Institute of Molecular Biotechnology, Genome Analysis, *
Peutenbergstrasse 11, D-07745 Jena, Germany,
* e-mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e-mail: shimizudmb-med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/
and
* GBF, Dept. of Genome Analysis,
Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e-mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
Innestrasse 73, D-14195 Berlin, Germany,
* e-mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.
Location/Qualifiers
1. 340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
<1. 125946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
/clone="P160L9, 5' partial"
/clone_lib="RPCIL3-5 PAC library"
/notes="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
/clone="P70124"
/clone_lib="RPCIL3-5 PAC library"
/notes="Accession No. AF064860"
224137..2340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPCIL3-5 PAC library"
/notes="Accession No. AF121782"
286628..2340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPCIL3-5 PAC library"
/notes="Accession No. AF121782"
286628..2340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPCIL3-5 PAC library"
/notes="Accession No. AF064857"
1088..1566
repeat_region

```

```

/note="L1M4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/note="THE1C"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..35157
/genes="SH3BGR"
join(<3473..3565,12649..12678,15331..15460,18739..19157)
/genes="SH3BGR"
join(3473..3565,12649..12678,15331..15426)
/partial
/genes="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CAB90445.1"
/db_xref="GI:7717370"
/translation="GSEKAEGGETEAKOEGSEDVGNLPDAQKNEBGETATETEE
IAMEGAEGBAEETAEGERPEDEDS"
3473..3565
/genes="SH3BGR"
/number=4
5542..5829
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="THE1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)

```

```

/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12649..12678
/genes="SH3BGR"
/number=5
complement(13115..13425)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(13598..13892)
/note="AluY"

Query Match      28.0%; Score 326.6; DB 9; Length 340000;
Best Local Similarity 97.4%; Pred. No. 4.9e-67; Indels 0; Gaps 0;
Matches 332; Conservative 0; Mismatches 9;

Qy 27 AAATAAGAGACCCACCGGTTCTGGGTCTGTGAATGAAGTCATAGAGGGCCCCCAAAATG 86
Db 268437 AATTGGCTACCTTCCAGGTTCTGGGTCTGTGAATGAAGTCATAGAGGGCCCCCAAAATG 268496
Qy 87 CAGAGTCTTGAGGGGCTCCAGGCTCGTTCACCTGCACCGTCTCCAGGGCTGGAAGC 146
Db 268497 CAGAGTCTTGAGGGGCTCCAGGCTCGTTCACCTGCACCGTCTCCAGGGCTGGAAGC 268556
Qy 147 TCATCATGTGGGCTCTCAGTGACATGGTGTGCTTAAGCGTCAGGCCCATCGAGGCCCATCA 206
Db 268557 TCATCATGTGGGCTCTCAGTGACATGGTGTGCTTAAGCGTCAGGCCCATCGAGGCCCATCA 268616
Qy 207 TCACCAATGACCGCTTCACCTCTCAGAGGTACGACGAGGGGGGGAACCTTCACTCGGAGA 266
Db 268617 TCACCAATGACCGCTTCACCTCTCAGAGGTACGACGAGGGGGGGAACCTTCACTCGGAGA 268676
Qy 267 TGATCATCCACAATGTGGAGCCAGTGATTGCGGGGAACATCAGATGCAGGCTCCAGAAC 326
Db 268677 TGATCATCCACAATGTGGAGCCAGTGATTGCGGGGAACATCAGATGCAGGCTCCAGAAC 268736
Qy 327 GTCGCTGCATGGATCTGCTTACCTTACCTCCAAAGTTATG 367
Db 268737 GTCGCTGCATGGATCTGCTTACCTTACCTCCAAAGTTATG 268777

RESULT 10
AC120145 149964 bp DNA linear HTG 09-MAR-2003
LOCUS Mus musculus clone RP24-216P24, WORKING DRAFT SEQUENCE, 9 unordered
DEFINITION pieces.
ACCESSION AC120145
VERSION AC120145.3 GI:28893697
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 149964)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP24-216P24

```

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 149964)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Bierre, N., Jones, C.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R.,  
Lander, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menes, L.,  
Mihova, T., Milena, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (03-MAY-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 149964)  
Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,  
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,  
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Faro, S.,  
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,  
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,  
Graham, L., Grand-Bierre, N., Hafez, N., Hagopian, D., Hagos, B.,  
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Lander, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,  
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
Meldrum, J., Menes, L., Mihova, T., Milena, V., Murphy, T., Naylor, J.,  
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
O'Neil, D., Olivier, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (09-MAR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 9, 2003 this sequence version replaced gi:28604025.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Location/Qualifiers  
1. .149964  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone\_lib="RP24-216P24"  
/clone\_lib="RP24-216P24"  
1. .582  
/note="assembly\_fragment"  
683. 1358  
/note="assembly\_fragment"  
1459. .3822  
/note="assembly\_fragment"  
3923. .8564  
/note="assembly\_fragment"  
8665. .21635  
/note="assembly\_fragment"  
21736. .40606  
/note="assembly\_fragment"  
40707. .60552  
/note="assembly\_fragment"  
60553. .90141  
/note="assembly\_fragment"  
90242. .149964  
/note="assembly\_fragment"  
clone\_end:17  
vector\_side:right

BASE COUNT 38568 a 33807 c 34845 g 41924 t 800 others  
ORIGIN

Quality coverage: 10.8 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

582: contig of 582 bp in length  
682: gap of 100 bp  
1358: contig of 676 bp in length  
1458: gap of 100 bp  
3822: contig of 2364 bp in length  
3922: gap of 100 bp  
8564: contig of 4642 bp in length  
8664: gap of 100 bp  
21635: contig of 12971 bp in length  
21735: gap of 100 bp  
21736: contig of 18871 bp in length  
40606: contig of 18871 bp in length  
40706: gap of 100 bp  
60552: contig of 19846 bp in length  
60552: gap of 100 bp  
90141: contig of 29489 bp in length  
90241: gap of 100 bp  
149964: contig of 59723 bp in length.

Quality Match 15.6%; Score 182.4; DB 2; Length 149964;  
Best Local Similarity 72.3%; Pred. No. 1.4e-32;  
Matches 237; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

37 CCACCCGGTTCGGGTCTGGTAATGAGTCTAGAGGCCGCCCAATGCAAGAGTCT 96  
DB 89205 CCTTCTAGCTTCGGATCCAGTATCATAGATCATAGAGGTCTCAGATGTACAGTCT 89264

97 GAAGGCTCCAGGCTCGCTTCACTGACCGCTCTCCAGGCTGGAAGCTCATGTG 156  
DB 89265 AAAGGACTCAGAGGCTCACTTCACTGACCGCTGACTACCGGCTTCTCATGTG 89324

157 GGCTCTCAGTGCATGGTGGTGTAAAGCGTCAAGCGCTAGGCCCATCATCAACCAATGA 216  
DB 89325 GACTCTTACCAAAATGGTGGTGTCTGAGTCTCACCACCCAGGACCATCATCAACCA 89384

TITLE  
JOURNAL  
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L25242  
Center clone name: 216\_P\_24  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 147153 bases at least Q40  
Consensus quality: 147612 bases at least Q30  
Consensus quality: 147899 bases at least Q20  
Insert size: 146000; agarose-fp  
Insert size: 149164; sum-of-contigs  
Quality coverage: 11.0 in Q20 bases; agarose-fp

Qy 217 CCGCTTACCTCTCAGAGTATACACAGGCGGGAACTTCACTCGGAGATGATCATCCA 276  
 |||||  
 Db 89385 CCGCTTACCTATGCCAGTTTACACAGCACTGACAGCTTCATCTCGAGATGATCATCCA 89444  
 |||||  
 Qy 277 CAATGTGGACCCAGTATTCGGGGACATCATCATGCGCTCCAGAAACAGTGGCCTGCA 336  
 |||||  
 Db 89445 TGATGTGACGCCAGTGTACTCGGATCCGTCGATGACGCTCGACAGCAGCATGGT 89504  
 |||||  
 Qy 337 TGGATCTGCTTACCTTACCTCCCAAGTT 364  
 |||||  
 Db 89505 TGGATCTGCTTCTCTCTCTCAGTGAAGT 89532  
 |||||

RESULT 11  
 AC120346 181510 bp DNA linear HTG 22-OCT-2002  
 LOCUS Mus musculus clone RP23-147E11, WORKING DRAFT SEQUENCE, 15 ordered  
 DEFINITION pieces.  
 AC120346  
 VERSION AC120346.4 GI:24211306  
 KEYWORDS HTG; HTGS\_PHRASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 181510)  
 Authors Birren, B., Nussbaum, C. and Lander, E.  
 TITLE Mus musculus, clone RP23-147E11  
 JOURNAL Unpublished  
 AUTHORS 2 (bases 1 to 181510)  
 REFERENCES Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
 Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
 Fato, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,  
 Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McSwan, P., McKernan, K., Meldrim, J., Meneus, L.,  
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 REFERENCE 3 (bases 1 to 181510)  
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
 Camarata, J., Chang, P., Chazaro, B., Choepel, Y., Collymore, A.,  
 Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
 Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,  
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,  
 Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,  
 Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,  
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (22-OCT-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Oct 22, 2002 this sequence version replaced gi:20514894.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information  
 Center project name: L20921  
 Center clone name: 147\_E\_11

----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 177801 bases at least Q40  
 Consensus quality: 179340 bases at least Q30  
 Consensus quality: 179851 bases at least Q20  
 Insert size: 18000; agarose-fp  
 Quality coverage: 11.3 in Q20 bases; agarose-fp  
 Quality coverage: 11.1 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\* 1 29098: contig of 29098 bp in length  
 \* 29099 29198: gap of 100 bp  
 \* 29199 29864: contig of 666 bp in length  
 \* 29865 29964: gap of 100 bp  
 \* 29965 30244: contig of 280 bp in length  
 \* 30245 30344: gap of 100 bp  
 \* 30345 30971: contig of 627 bp in length  
 \* 30972 31071: gap of 100 bp  
 \* 31072 32648: contig of 1577 bp in length  
 \* 32649 32748: gap of 100 bp  
 \* 32749 37186: contig of 4438 bp in length  
 \* 37187 37286: gap of 100 bp  
 \* 37287 42423: contig of 6957 bp in length  
 \* 42424 44343: gap of 100 bp  
 \* 44344 54253: contig of 9910 bp in length  
 \* 54254 54353: gap of 100 bp  
 \* 54354 70476: contig of 16123 bp in length  
 \* 70477 70576: gap of 100 bp  
 \* 70577 86055: contig of 15479 bp in length  
 \* 86056 86155: gap of 100 bp  
 \* 86156 104095: contig of 17940 bp in length  
 \* 104096 104195: gap of 100 bp  
 \* 104196 123740: contig of 19545 bp in length  
 \* 123741 123840: gap of 100 bp  
 \* 123841 153195: contig of 29355 bp in length  
 \* 153196 153295: gap of 100 bp  
 \* 153296 180206: contig of 26911 bp in length  
 \* 180207 180306: gap of 100 bp  
 \* 180307 181510: contig of 1204 bp in length.

Location/Qualifiers  
 1. 181510  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-147E11"  
 /clone\_lib="RP23-147E11" Female Mouse BAC"

FEATURES  
 source

```
misc_feature 1. 29098
/note="assembly_fragment
clone_end:Sp6
vector_side:left"
misc_feature 29199..29864
/note="assembly_fragment"
misc_feature 29965..30244
/note="assembly_fragment"
misc_feature 30345..30971
/note="assembly_fragment"
misc_feature 31072..32648
/note="assembly_fragment"
misc_feature 32749..37186
/note="assembly_fragment"
misc_feature 37287..44243
/note="assembly_fragment"
misc_feature 44344..54253
/note="assembly_fragment"
misc_feature 54354..70476
/note="assembly_fragment"
misc_feature 70577..86055
/note="assembly_fragment"
misc_feature 86156..104095
/note="assembly_fragment"
misc_feature 104196..123740
/note="assembly_fragment"
misc_feature 123841..153195
/note="assembly_fragment"
misc_feature 153296..180206
/note="assembly_fragment"
misc_feature 180307..181510
/note="assembly_fragment
clone_end:T7
vector_side:right"

BASE COUNT 49236 a 40873 c 40708 g 49292 t 1401 others
ORIGIN

Query Match 15.6%; Score 182.4; DB 2; Length 181510;
Best Local Similarity 72.3%; Pred. No. 1.4e-32;
Matches 237; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 37 CCCACCGGTTCTGGTCTGTGTAATGAAGTCATAGAAGGCCGCCCAAAATGCAAGAGTCTT 96
|||
Db 152357 CTTCTAGCTTCGGATCCAGTTATCAGATCATAGAAGTCTCAGATGTAACAGTCTT 152416
|||

QY 97 GAAGGCTCCAGGCTCGTTCACTGCACCGTCTCCAGGCTGGAAGCTCATCATGTG 156
|||
Db 152417 AAAGGACTCAGAGGCTCACTTCACTGCACCGTCACTCAGGCTGGAAGCTTCTCATGTG 152476
|||

QY 157 GGCTCTCAGTGACATGGTGGTGTCTAAGGCTCAGGCGCCCATGAGCCCATCACCACATGA 216
|||
Db 152477 GACTCTTACCAATAGTGTGTGCTGAGTCTCACCCACCAAGACCCATCATCACCAAA 152536
|||

QY 217 CGCTTCACCTCTCAGAGGTAGCAGCAGGCGGGAACTTCACTCGGAGATGATCATCCA 276
|||
Db 152537 CGCTTCACCTATGCCAGTTACACAGCAGCTGACAGCTTCACTCGGAGTTGATCATCCA 152596
|||

QY 277 CAATGTGGAGCCAGTGATTCGGGGAACATCAGATGACGCTCCAGACAGTCGCTGCA 336
|||
Db 152597 TGATGTGACCCAGTACGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 152656
|||

QY 337 TGATCTGCTTACCTTACCGTCAAGTT 364
|||
Db 152657 TGGATCTGCTTCTCTCATGTGCAAGT 152684
|||

RESULT 12
AC020851
LOCUS Mus musculus clone RP21-467L12, WORKING DRAFT SEQUENCE, 33
DEFINITION unorderd pieces.
ACCESSION AC020851
VERSION AC020851.2 GI:9211212
```

HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
Mus musculus (house mouse)  
Mus musculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 196900)  
DOE Joint Genome Institute.  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 196900)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 15, 2000 this sequence version replaced gi:6686457.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1437240  
Center clone name: RPI-21\_467L12  
-----  
Summary Statistics  
Consensus quality: 164618 bases at least Q40  
Consensus quality: 180887 bases at least Q30  
Consensus quality: 183777 bases at least Q20  
Estimated insert size: 147000; pulse field gel estimation  
Estimated insert size: 193700; sum-of-contigs estimation  
Quality coverage: 5.67 in Q20 bases; pulse field gel estimation  
Quality coverage: 4.3 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 33 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1541: contig of 1541 bp in length  
\* 1542 1641: gap of unknown length  
\* 1642 3492: contig of 1851 bp in length  
\* 3493 3592: gap of unknown length  
\* 3593 4738: contig of 1146 bp in length  
\* 4739 4838: gap of unknown length  
\* 4839 7248: contig of 2410 bp in length  
\* 7249 7349: gap of unknown length  
\* 7349 9179: contig of 1830 bp in length  
\* 9179 9279: gap of unknown length  
\* 9279 11493: contig of 2215 bp in length  
\* 11494 11593: gap of unknown length  
\* 11594 13607: contig of 2014 bp in length  
\* 13608 13707: gap of unknown length  
\* 13708 15946: contig of 2239 bp in length  
\* 15947 16047: gap of unknown length  
\* 16047 18392: contig of 2346 bp in length  
\* 18393 18492: gap of unknown length  
\* 18493 21231: contig of 2739 bp in length  
\* 21232 21331: gap of unknown length  
\* 21332 24662: contig of 3331 bp in length  
\* 24663 24762: gap of unknown length  
\* 24763 27140: contig of 2378 bp in length  
\* 27141 27240: gap of unknown length  
\* 27241 29950: contig of 2710 bp in length  
\* 29951 30050: gap of unknown length  
\* 30051 32245: contig of 2195 bp in length  
\* 32246 32345: gap of unknown length  
\* 32346 35688: contig of 3343 bp in length  
\* 35689 35788: gap of unknown length  
\* 35789 40387: contig of 4599 bp in length  
\* 40388 40487: gap of unknown length  
\* 40488 43580: contig of 3093 bp in length

KEYWORDS  
SOURCE ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

```

* 43581 43680: gap of unknown length
* 43681 46437: contig of 2757 bp in length
* 46438 46537: gap of unknown length
* 46538 50184: contig of 3647 bp in length
* 50185 50284: gap of unknown length
* 50285 54572: contig of 4288 bp in length
* 54573 54672: gap of unknown length
* 54673 60227: contig of 5555 bp in length
* 60228 60328: gap of unknown length
* 60328 68024: contig of 7697 bp in length
* 68025 68124: gap of unknown length
* 68125 78510: contig of 10486 bp in length
* 78511 78710: gap of unknown length
* 78711 88151: contig of 9441 bp in length
* 88152 88251: gap of unknown length
* 88252 96257: contig of 8006 bp in length
* 96258 96358: gap of unknown length
* 96358 106228: contig of 9870 bp in length
* 106228 106328: gap of unknown length
* 106328 115795: contig of 9468 bp in length
* 115796 115895: gap of unknown length
* 115896 126363: contig of 10468 bp in length
* 126364 126463: gap of unknown length
* 126464 137189: contig of 10726 bp in length
* 137190 137289: gap of unknown length
* 137290 146378: contig of 9089 bp in length
* 146379 146478: gap of unknown length
* 146479 161686: contig of 15208 bp in length
* 161687 161786: gap of unknown length
* 161787 176149: contig of 14363 bp in length
* 176150 176249: gap of unknown length
* 176250 196900: contig of 20651 bp in length.

FEATURES             Location/Qualifiers
     source            1..196900
     organism="Mus musculus"
     mol_type="genomic DNA"
     db_xref="taxon:10090"
     clone="RP21-467L12"
     /clone_lib="RP21 mouse PAC library 21"
BASE COUNT      51456 a 43823 c 44608 g 53666 t 3347 others
ORIGIN

Query Match      15.6%; Score 182.4; DB 2; Length 196900;
Best Local Similarity 72.3%; Pred. No. 1.4e-32;
Matches 237; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 37 CCCACCCGGTCTGGGCTGTAATGAGTATAGAGGCCCCCAAAATGCAAGAGTCCCT 96
Db 67439 CCTTCTAGCTCCGATCCAGTTATCAGATCATAGAGGCTCTCAGAATGTAAACAGTCT 67498

QY 97 GAAGGCTCCCGGCTGCTTCACTGACCGCTCCCGAGGCTGGAAGCTCATCATGTG 156
Db 67499 AAAGGACTAGAGGCTCACTTCACTGACCGTCTGAGTCTGAGGCTGGAAGCTTCTCATGTG 67558

QY 157 GGCTCTCAGTGCATGGTGGTCTAAGCGTCTAGGCGCCATGAGGCGCCATCATCCCAATGA 216
Db 67559 GACTCTTAACCAAAATGGTGGTCTGAGTCTCACCACCAAGGACCATCATCCCAACAA 67618

QY 217 CGCTTTCACCTCTCAGAGGTAGACAGCGGGGAGCTTCACTCGGAGATCATCATCA 276
Db 67619 CCGCTTCACTTATGCCAGTTTAAACAGACTGACAGCTTCACTCGGAGTTCATCATCA 67678

QY 277 CAATGTGAGCCAGTGTATCGGGGAACATCAGATCAGCCCTCCAGAACAGTCCGCTGCA 336
Db 67679 TGATGTGACGCCAGTGAAGTCTCGGATTCGTCGAATGACGCTGCAAGACAGCCATGGGTT 67738

QY 337 TGATGTGCTTACCTTACCGTCCAAGTT 364
Db 67739 TGGATCTGCTTCTCTCAGTGAAGGT 67766

RESULT 13
AX380408

```

```

LOCUS      AX380408      754 bp      DNA      linear      PAT 18-MAR-2002
DEFINITION Sequence 13 from Patent WO0200710.
ACCESSION  AX380408
VERSION    AX380408.1  GI:19575338
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1
AUTHORS   Melcher,A.A., Samtiento,U.M., Schultz,H.U. and Chute,H.T.
TITLE     B7-like molecules and uses thereof
JOURNAL   Patent: WO 0200710-A 13 03-JAN-2002;
          Angen, Inc. (US)
FEATURES             Location/Qualifiers
     source            1..754
     organism="Mus musculus"
     mol_type="genomic DNA"
     db_xref="taxon:10090"
     53..724
     /notes="unlabeled protein product"
     /codon_start=1
     /protein_id="CAD28222.1"
     /db_xref="GI:19575338"
     /translation="MEGSRDVLAVLVLAQLTAGSSVQIIEGPQNTVTLKDSBHP
     NCTVTHGWLKMLTLMQVVLSTTQGPITNNRFTYASNTDSFISLIHVDQPS
     DSGSVQCSLQNSHGFGSFLVSQSTYQNEIRKSANMTNKADPTKLSKNGKYGYS
     SDEAKAAQTASLPPKSAEVSLEPKRSSSLPYQLNKHQPGPATHPRVSPFIASPKVR
     NVTLV"
BASE COUNT      220 a 191 c 175 g 168 t
ORIGIN

Query Match      15.6%; Score 181.8; DB 6; Length 754;
Best Local Similarity 72.9%; Pred. No. 1.4e-32;
Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 42 CCGGTTCTGGGCTGCTGAATGAGTATAGAGTATAGAGGCCCCCAAAATGCAAGAGTCCCTCAAGG 101
Db 108 CAGCTTCGGATCCAGTTATCAGATCATAGAGGCTCTCAGAATGTAAACAGTCTTAAAGG 167

QY 102 GCTCCAGGCTCGCTTCACTGACCGCTCTCCAGGGCTGGAGCTCATCATGTGGGCTC 161
Db 168 ACTCAGAGGCTCACTTCACTGACCGCTGACTCAGCGCTGGAAGCTTCTCATGTGGACTC 227

QY 162 TCAGTGACATGGTGGTCTAAGCGTCTAGGCGCCATGAGGCCCATCATCAACATGACCGCT 221
Db 228 TTAACCAAAATGGTGGTCTGAGTCTCACCACCAAGGACCCCATCATCAACCAACCGCT 287

QY 222 TCACCTCTCAGAGGTACACCGAGGGGAACTTTCACCTCGGAGATGATCATCCCAATG 281
Db 288 TCACCTATGCCAGTTACACAGCACTGACAGCTTCACTCGGAGTTGATCATCATGATG 347

QY 282 TGGAGCCCATGATTTGGGGAACATCAGATGAGCGCTTCAGAACAGTCTCGCTCGATGAT 341
Db 348 TGCAGCCCATGATCTGGGATCCGTCGAATGAGCGCTGAGAGGACAGCCATCGGCTTTGGAT 407

QY 342 CTGCTTACCTTACCGTCCAAAG 362
Db 408 CTGCTTCTCTCAGTGCAG 428

RESULT 14
AX380406
LOCUS      AX380406      895 bp      DNA      linear      PAT 18-MAR-2002
DEFINITION Sequence 11 from Patent WO0200710.
ACCESSION  AX380406
VERSION    AX380406.1  GI:19575336
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1

```

AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.  
 TITLE B7-like molecules and uses thereof  
 JOURNAL Patent: WO 0200710-A 11 03-JAN-2002;  
 Angen, Inc. (US)

FEATURES  
 source 1. .895  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 53. .865  
 /notes="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD28221.1"  
 /db\_xref="GI:19575337"  
 /translation="MEGSRWDVLAVLILVLAQLTASGSSQIIEGPNVTVLKDEAHF  
 NCTVTHGKMLNQMVLSTTQGLITNNRFTYASINSTDSFISLIIHDVQPS  
 DSGVQCSLQNSHGFGSIVSDSISGEGPALFWAILLAVAFSLIILIVLII  
 FCCCAARKEESTYQNIKRSANRNTKADPTKLGKNGYSSDEAKAAQTAS  
 LPPKSAEVLPEKSSSLPYQLNKHQPGPATHPRVPSFDIASPQKRVNLTUV"

BASE COUNT 250 a 223 c 212 g 210 t  
 ORIGIN

Query Match 15.6%; Score 181.8; DB 6; Length 895;  
 Best Local Similarity 72.94; Pred. No. 1.4e-32;  
 Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 42 CCGGTTCTGGGTCTGGTAAGTTCATAGAAGGCCCAAAATGCAAGAGTCCTGAAGG 101  
 |||||  
 Db 108 CAGCTTCGGATCCAGTTATCAGTATCAGTAAGGTCCTCAGATGTAACAGTCCTTAAGG 167  
 |||||  
 QY 102 GTCGCCAGGTCGTTCACTGACCGTCTCCAGGCGTGGAGCTCATCATGTGGGCTC 161  
 |||||  
 Db 168 ACTCAGAGGCTCACTTCAACTGACCGTGACTCAGCGCTGGAGCTTCTCATGTGGACTC 227  
 |||||  
 QY 162 TCAGTGACATGGTGTCTAAGGCTCAGGCGCCATGAGCCATCATCAACCAATGACCGCT 221  
 |||||  
 Db 228 TTACCAAAATGGTGTCTGAGTCTCACCACCAAGAGCCCATCATCAACCAACCGCT 287  
 |||||  
 QY 222 TCACCTCTCAGAGTACGACACGCGGGGAACTTCACTCGGAGATGATCATCAACAATG 281  
 |||||  
 Db 288 TCACCTATGCGAGTTACACACGACTGACAGCTTCATCTCGGAGTTGATCATCATGATG 347  
 |||||  
 QY 282 TGGAGCCCACTGATTCGGGGAATCATCAGATCGAGCTCCAGAACAGTCGCGCTCATGGAT 341  
 |||||  
 Db 348 TGCAGCCCACTGATTCGGGATCCGTCGATCGAGCTTCGAGAACAGCATGCGGTTGGAT 407  
 |||||  
 QY 342 CTGCTTACCTTACCGTCCAAAG 362  
 |||||  
 Db 408 CTGCGTTCTCTCAGTGCAG 428  
 |||||

RESULT 15  
 BC004806 1556 bp mRNA linear ROD 16-APR-2003  
 LOCUS Mus musculus RIKEN cDNA 2010003D20 gene, mRNA (cDNA clone MGC:7960  
 IMAGE:3584645), complete cds.

ACCESSION BC004806  
 VERSION BC004806.1 GI:13435932  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1556)

REFERENCE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22382257  
 12477932  
 2 (bases 1 to 1556)  
 Strausberg, R.  
 Direct Submission  
 Submitted (21-MAR-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IRAK Plate: 10 Row: m Column: 23  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

## FEATURES

Location/Qualifiers  
 1. .1556  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:7960 IMAGE:3584645"  
 /tissue\_type="Mammary tumor. Metallothionien-TGF alpha  
 model. 10 month old virgin mouse. Taken by biopsy."  
 /clone\_lib="NCI CGAP\_Mam1"  
 /lab\_host="DHI0B"  
 /note="Vector: pCMV-SPORT6"  
 1. .1556  
 /gene="2010003D20Rik"  
 /db\_xref="LocusID:72058"  
 /db\_xref="MGI:1919308"  
 55. .867  
 /codon\_start=1  
 /product="RIKEN cDNA 2010003D20"  
 /protein\_id="AAH04806.1"  
 /db\_xref="GI:13435933"  
 /db\_xref="LocusID:72058"  
 /translation="MEGSRWDVLAVLILVLAQLTASGSSQIIEGPNVTVLKDEAHF  
 NCTVTHGKMLNQMVLSTTQGLITNNRFTYASINSTDSFISLIIHDVQPS  
 DSGVQCSLQNSHGFGSIVSDSISGEGPALFWAILLAVAFSLIILIVLII  
 FCCCAARKEESTYQNIKRSANRNTKADPTKLGKNGYSSDEAKAAQTAS  
 LPPKSAEVLPEKSSSLPYQLNKHQPGPATHPRVPSFDIASPQKRVNLTUV"

## gene

## CDS

## BASE COUNT

## Query Match

15.6%; Score 181.8; DB 10; Length 1556;

Best Local Similarity 72.9%; Pred. No. 1.5e-32;  
 Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY	42	COGGTTCTGGGTCGTGAATGAAGTCATAGAAGGCCCCCAAAATGCAAGAGTCCTCAAGG	101
Db	110	CAGCTTCGGATCCAGTTATCAGATCATAGAAGGTCCTCAGAAATGTAAACAGTCCCTAAAGG	169
QY	102	GCTCCAGGTCGCTTCAAACCTCACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTC	161
Db	170	ACTCAGAGGCTCACITCAACTGCACCGTGACTCACGGCTGGAAGCTTCTCATGTGGACTC	229
QY	162	TCAGTGACATGGTGTGCTTAAGCGTCAGGCCCATGAGGCCCATCATCAACATGACCGCT	221
Db	230	TTAACCAAAATGGTGTGCTGAGTCTCACCCACCAAGGACCCCATCATCAACCAACACCGCT	289
QY	222	TCACCTCTCAGAGGTACGACCCAGGGCGGAACTTCACCTCGGAGATGATCATCCACAATG	281
Db	290	TCACCTATGCCAGTTACAAACAGCACTGACAGCTTCATCTCGAGTTGATCATCCATGATG	349
QY	282	TGGAGCCCATGATTCGGGGAAACATCAGATGAGCCTCCAGAACACAGTCGCTGCAATGGAT	341
Db	350	TGCAGCCCATGACTCGGGATCCGTGCAATGAGCCTGCAAGACAGCCATGGGTTTGGAT	409
QY	342	CTGCTTACTTACCGTCCAAG	362
Db	410	CTGCTTCTCTCAGTGCAAG	430

Search completed: November 12, 2003, 20:56:33  
 Job time : 4520.2 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:39:35 ; Search time 345.543 Seconds

(without alignments)  
9124.610 Million cell updates/sec

Title: US-09-729-264-3

Perfect score: 1168

Sequence: 1 agtgatcatggggcaggag.....gtaatacaactgtagtatag 1168

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1168	100.0	1168	24	ABK13029 DNA encoding human
2	1122.4	96.1	1175	24	ABK13028 DNA encoding human
3	1119	95.8	1240	24	ABK13030 DNA encoding human
4	1102.8	94.4	1139	24	ABK13031 DNA encoding human
5	777.6	66.6	1392	23	AAS92356 DNA encoding novel
6	378.4	32.4	1195	24	ABK13032 DNA encoding mouse
7	196.8	16.8	401	22	AA136582 Probe #5268 used t
8	181.8	15.6	754	24	ABK13034 DNA encoding mouse

9	181.8	15.6	895	24	ABK13033	DNA encoding mouse
10	136.2	11.7	398	25	ABX55016	Bovine EST associa
c 11	59.4	4.6	277	21	ABK10239	Tri-nucleotide repe
12	52.8	4.5	7029	23	ABL29756	Drosophila melanog
c 13	51.8	4.4	389	21	ABK10248	Tri-nucleotide rep
c 14	51.8	4.4	397	20	AAK98991	Spinocerebellar at
c 15	51.8	4.4	403	21	ABK10240	Tri-nucleotide repe
c 16	49.4	4.2	46954	23	ABK16830	Drosophila melanog
c 17	48.4	4.1	3217	10	AAK91578	Rat androgen recep
c 18	48.4	4.1	3217	12	AAQ12002	Full-length rat an
c 19	48.4	4.1	4180	10	AAK91773	Rat androgen recep
c 20	47.8	4.1	381	23	ABV54466	Human prostate exp
c 21	47.6	4.1	2733	20	AAK84442	Mouse brain CNG-1
c 22	47.4	4.1	3894	23	ABK03353	Drosophila melanog
c 23	47.4	4.1	22341	23	ABK03352	Drosophila melanog
c 24	46.6	4.0	2333	22	AAK75339	Human TGF-beta rec
c 25	46.4	4.0	1983	23	ABL24483	Drosophila melanog
c 26	46.4	4.0	4044	23	ABL24482	Drosophila melanog
c 27	46	3.9	3577	23	ABK16536	Canine cDNA encodi
c 28	45.8	3.9	2682	23	ABL16097	Drosophila melanog
c 29	45.8	3.9	8821	23	ABL16096	Drosophila melanog
c 30	45.6	3.9	1509	23	ABL24571	Drosophila melanog
c 31	45.6	3.9	3884	23	ABL24570	Drosophila melanog
c 32	45.4	3.9	3042	23	ABL28445	Drosophila melanog
c 33	45.4	3.9	3135	23	ABL25653	Drosophila melanog
c 34	45.4	3.9	5215	23	ABL28444	Drosophila melanog
c 35	45.4	3.9	5447	23	ABL25652	Drosophila melanog
c 36	45.2	3.9	5397	23	ABL29757	Drosophila melanog
c 37	45	3.9	462	23	ABL26941	Drosophila melanog
c 38	45	3.9	2115	23	ABL26940	Drosophila melanog
c 39	45	3.9	2260	23	ABL03191	Drosophila melanog
c 40	45	3.9	3210	23	ABL08977	Drosophila melanog
c 41	45	3.9	3624	23	ABL03190	Drosophila melanog
c 42	45	3.9	5975	23	ABL08976	Drosophila melanog
c 43	45	3.9	6541	23	ABL04172	Drosophila melanog
c 44	44.8	3.8	254	21	AAZ45758	Polymorphic region
c 45	44.8	3.8	1024	25	ABZ83769	Toxicologically re

## ALIGNMENTS

### RESULT 1

ABK13029

ID ABK13029 standard; cDNA; 1168 BP.

AC ABK13029;

DT 23-APR-2002 (first entry)

XX DNA encoding human B7-like protein, B7-L\_h2.

DE Human; B7-like protein; B7-L; antiinfectivity; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antifungal; antiallergic;  
KW antidiabetic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 8..1168

FT /tag= a

FT /product= "B7-like protein, B7-L\_h2"

XX WO200200710-A2.

PD 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US20719.

PF

XX 28-JUN-2000; 2000US-214512P.  
PR 28-NOV-2000; 2000US-0729264.  
XX (AMGE-) AMGEN INC.  
XX Weicher AA, Sarmiento UM, Schultz HJ, Chute HT;  
PI WPI; 2002-130881/17.  
XX P-PSDB; AAU75541.  
XX  
PT New B7-like polypeptides, polynucleotides and their modulators, useful  
PT for diagnosing, preventing and treating reproductive, immune and  
PT proliferative disorders, e.g. cancer and arteriosclerosis -  
XX Claim 1; Fig 2; 135pp; English.  
XX  
CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
CC useful for treating B7-like polypeptide-related disease, disorders or  
CC conditions including reproductive disorders (e.g. infertility,  
CC miscarriage, preterm labour and delivery and endometriosis) and  
CC proliferative disorders. Antibodies, soluble proteins comprising  
CC extracellular domains and other regulators of B7-L polypeptides are  
CC useful for enhancing the immune response to tumours. (I) plays a role in  
CC growth and maintenance of cancer cells based on the observation of  
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
CC polypeptide. Hence modulators of (I) are useful for the treatment of  
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
CC in allograft transplantation, graft versus host disease, T-cell  
CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
CC molecules are useful for alleviating the symptoms associated with  
CC diseases involving chronic immune cell dysfunction or to treat  
CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
CC purpura and psoriasis, chronic inflammatory disease such as  
CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
CC are also useful as immunosuppressive agents for bone marrow and organ  
CC transplantation or to prolong graft survival. B7-L molecules are also  
CC useful for diagnosis and treatment of diseases involving abnormal cell  
CC proliferation, including arteriosclerosis and vascular restenosis.  
CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
CC shock syndrome or allo sensitisation due to blood transfusions, and for  
CC treatment of allergy, asthma and hypersensitivity reactions,  
CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
CC (extrinsic alveolitis) vasculopathies, coeliac disease, anaemia, and  
CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
CC lymphoproliferative disorders such as multiple myeloma. The present  
CC sequence represents the coding sequence of human B7-L<sub>h2</sub>.  
XX  
SQ Sequence 1168 BP; 299 A; 302 C; 288 G; 279 T; 0 other;

Db	181		AAGCGTCAGGCCCATGGAGCCCATCATCAACAAATGACCGCTTACCTCTTCAGAGGTACGA	240
Qy	241		CCAGGCGGGAACTTTCACCTCGGAGATGATCATCCACAATGTGTGAGCCAGTGTGGG	300
Db	241		CCAGGCGGGAACTTTCACCTCGGAGATGATCATCCACAATGTGTGAGCCAGTGTGGG	300
Qy	301		GAACATCAGATGCAGCCTCCAGAAACAGTCGCTGCATGGATCTCTTACCTTACCGTCCA	360
Db	301		GAACATCAGATGCAGCCTCCAGAAACAGTCGCTGCATGGATCTCTTACCTTACCGTCCA	360
Qy	361		AGTTATGGGAGAGCTGTTTCAATCCAGTGTAAATCTTTGTAGTCTCGAGAAATGAACCTTG	420
Db	361		AGTTATGGGAGAGCTGTTTCAATCCAGTGTAAATCTTTGTAGTCTCGAGAAATGAACCTTG	420
Qy	421		TGAAGTTACTTGTCTACCTCCACTCGACTCGCTCCCGGATATTTCTCGGAGCTCGG	480
Db	421		TGAAGTTACTTGTCTACCTCCACTCGACTCGCTCCCGGATATTTCTCGGAGCTCGG	480
Qy	481		TCTCTCGTTCAGCCATTCAAGCTATTATTTGTTCCGAGCCAGCCAGTTCCTCAAGTGC	540
Db	481		TCTCTCGTTCAGCCATTCAAGCTATTATTTGTTCCGAGCCAGCCAGTTCCTCAAGTGC	540
Qy	541		AGTGAGCATCTGCTCTGACCCACAGAGCAATGGGACTTTGACTTGGCTGGCTACCTG	600
Db	541		AGTGAGCATCTGCTCTGACCCACAGAGCAATGGGACTTTGACTTGGCTGGCTACCTG	600
Qy	601		GAAGAGCTTGAAGCCCGCAGTCTGCAACTGTAAATCTCACTGTGATTCGGTGTCCCA	660
Db	601		GAAGAGCTTGAAGCCCGCAGTCTGCAACTGTAAATCTCACTGTGATTCGGTGTCCCA	660
Qy	661		AGACATGGAGGTGGTATTATTAATTCAGGTGTATTAATCAAGTTTACCGAGTTTGGT	720
Db	661		AGACATGGAGGTGGTATTATTAATTCAGGTGTATTAATCAAGTTTACCGAGTTTGGT	720
Qy	721		TTCAATTCCTACTTGGGGGCAAGTTGACCTGGACTAGCAGGACCATGCTCTGAGCC	780
Db	721		TTCAATTCCTACTTGGGGGCAAGTTGACCTGGACTAGCAGGACCATGCTCTGAGCC	780
Qy	781		GACGTGTACTTCTTCAATACGCTCTGCTGCTGCGCGCGCTGCTTGTGTGGCTGCAACTG	840
Db	781		GACGTGTACTTCTTCAATACGCTCTGCTGCTGCGCGCGCTGCTTGTGTGGCTGCAACTG	840
Qy	841		CTGCTGCGGTGTTGTTTCTGCTGTAGAGAAAAGAGGATTCGTATTCAATTTCAAAA	900
Db	841		CTGCTGCGGTGTTGTTTCTGCTGTAGAGAAAAGAGGATTCGTATTCAATTTCAAAA	900
Qy	901		GAAATCTGAAAAGAGAGAACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAATCTC	960
Db	901		GAAATCTGAAAAGAGAGAACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAATCTC	960
Qy	961		CGGCTACAATTCAGATGAACAAAAAGACACAGACACCGCTTCTCTCCCTCCCAATCCTG	1020
Db	961		CGGCTACAATTCAGATGAACAAAAAGACACAGACACCGCTTCTCTCCCTCCCAATCCTG	1020
Qy	1021		TGAATCAGTGTATCTGAAACAAAGAAACAGTAGTGTGGCCCTCTCTACACGCGGCTGA	1080
Db	1021		TGAATCAGTGTATCTGAAACAAAGAAACAGTAGTGTGGCCCTCTCTACACGCGGCTGA	1080
Qy	1081		TCAACGTCCACCCAGGCGCAGCAAGTCAATCAACAGGCTTCTTTTAATCTGGCCAGTCTCTGA	1140
Db	1081		TCAACGTCCACCCAGGCGCAGCAAGTCAATCAACAGGCTTCTTTTAATCTGGCCAGTCTCTGA	1140
Qy	1141		GAAGGTCAAGTAATACTGTAGTATAG 1168	
Db	1141		GAAGGTCAAGTAATACTGTAGTATAG 1168	

RESULT 2  
ABK13028  
ID ABK13028 standard; cDNA; 1175 BP.  
XX AC ABK13028;

XX 23-APR-2002 (first entry)  
XX DNA encoding human B7-like protein, B7-L\_h1.  
XX Human; B7-like protein; B7-L; antinfertility; gynaecological;  
XX anti-tumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
XX antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
XX antidiabetic; haemostatic; antithyroid; antituber; antiallergic;  
XX antiaschmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
XX reproductive disorder; graft versus host disease; autoimmune disease;  
XX toxic shock syndrome; allergy; nephropathy; skin disorder;  
XX endocrinopathy; lymphoproliferative disorder; gene; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 27..1175  
XX /\*tag= a  
XX /product= "B7-like protein, B7-L\_h1"  
XX  
XX WO200200710-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 28-JUN-2001; 2001WO-US20719.  
XX  
XX 28-JUN-2000; 2000US-214512P.  
XX 28-NOV-2000; 2000US-0729264.  
XX (AMGE-) AMGEN INC.  
XX  
XX Weicher AA, Sarmiento UM, Schultz HJ, Chute HT;  
XX WPI; 2002-130881/17.  
XX P-PSDB; AAU75540.  
XX  
XX New B7-like polypeptides, polynucleotides and their modulators, useful  
XX for diagnosing, preventing and treating reproductive, immune and  
XX proliferative disorders, e.g. cancer and arteriosclerosis -  
XX Claim 1; Fig 1; 135pp; English.  
XX  
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
XX useful for treating B7-like polypeptide-related disease, disorders or  
XX conditions including reproductive disorders (e.g. infertility,  
XX miscarriage, preterm labour and delivery and endometriosis) and  
XX proliferative disorders. Antibodies, soluble proteins comprising  
XX extracellular domains and other regulators of B7-L polypeptides are  
XX useful for enhancing the immune response to tumours. (I) plays a role in  
XX growth and maintenance of cancer cells based on the observation of  
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
XX polypeptide. Hence modulators of (I) are useful for the treatment of  
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
XX in allograft transplantation, graft versus host disease, T-cell  
XX dependent B-cell mediated diseases and autoimmune diseases. B7-L  
XX molecules are useful for alleviating the symptoms associated with  
XX diseases involving chronic immune cell dysfunction or to treat  
XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
XX purpura and psoriasis, chronic inflammatory disease such as  
XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
XX are also useful as immunosuppressive agents for bone marrow and organ  
XX transplantation or to prolong graft survival. B7-L molecules are also  
XX useful for diagnosis and treatment of diseases involving abnormal cell  
XX proliferation, including arteriosclerosis and vascular restenosis.  
XX Antagonists of B7-L polypeptides are useful for alleviation of toxic  
XX shock syndrome or allo sensitisation due to blood transfusions, and for  
XX treatment of allergy, asthma and hypersensitivity reactions.

CC	nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
CC	pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
CC	(extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and
CC	thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
CC	lymphoproliferative disorders such as multiple myeloma. The present
CC	sequence represents the coding sequence of human B7-L_h1.
XX	
SQ	Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 other;
	Query Match 96.1%; Score 1122.4; DB 24; Length 1175;
	Best Local Similarity 99.5%; Pred. No. 6.4e-303; Indels 0; Gaps 0;
	Matches 1126; Conservative 0; Mismatches 6;
QY	37 CCCACCCGGTCTCGGTCTGGTGAATGAAGTCAATGAAGGCCCCCAAAATGCAAGAGTCTCT 96
DB	44 CCTCCACGGTCTCGGTCTGGTGAATGAAGTCAATGAAGGCCCCCAAAATGCAAGAGTCTCT 103
QY	97 GAAGGGCTCCACGGCTCGCTTCACTGACACGCTCTCCAGGGCTGGAAGCTCATCATGTG 156
DB	104 GAAGGGCTCCACGGCTCGCTTCACTGACACGCTCTCCAGGGCTGGAAGCTCATCATGTG 163
QY	157 GGCTCTCAGTGACATGCTGGTGTCTAAGCGTCAGGCCCATGGAGCCCATCATCAACAATGA 216
DB	164 GGCTCTCAGTGACATGCTGGTGTCTAAGCGTCAGGCCCATGGAGCCCATCATCAACAATGA 223
QY	217 CGCTTCACTCTCAGAGGTACGACAGGGGGAACTTCACTCGGAGATGATCATCCA 276
DB	224 CGCTTCACTCTCAGAGGTACGACAGGGGGAACTTCACTCGGAGATGATCATCCA 283
QY	277 CAATGTGGAGCCCGAGTTCGGGGACATCAGATGAGCCTCCAGAACAGTCGGCTGCA 336
DB	284 CAATGTGGAGCCCGAGTTCGGGGACATCAGATGAGCCTCCAGAACAGTCGGCTGCA 343
QY	337 TGGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCATTCCAGTGTAACTCT 396
DB	344 TGGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCATTCCAGTGTAACTCT 403
QY	397 TGTAGTCGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGACCTGGCT 456
DB	404 TGTAGTCGCTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCTCAGCTGACCTGGCT 463
QY	457 CCCGGATATTTCCCTGGGAGCTCGGTCTCTCGTCAGCCATTCAGCTATTAATTTTGTTC 516
DB	464 CCCGGATATTTCCCTGGGAGCTCGGTCTCTCGTCAGCCATTCAGCTATTAATTTTGTTC 523
QY	517 GGAGCCAGGACCTTCAAAGTCAGTGAGCATCTGCTGCTGACCCACAGAGCAATGG 576
DB	524 GGAGCCAGGACCTTCAAAGTCAGTGAGCATCTGCTGCTGACCCACAGAGCAATGG 583
QY	577 GACTTGTGCTTGGCTGCTACCTCGAAGAGCCTGAAAGCCCGCAAGTCTGCAACTGTAAA 636
DB	584 GACTTGTGCTTGGCTGCTACCTCGAAGAGCCTGAAAGCCCGCAAGTCTGCAACTGTAAA 643
QY	637 TCTCAGTGTGATTCGGGTGCTCCCAAGACACTGGAGGTGTTAATTAATTCAGGTGTTAT 696
DB	644 TCTCAGTGTGATTCGGGTGCTCCCAAGACACTGGAGGTGTTAATTAATTCAGGTGTTAT 703
QY	697 ATCAAGTTTACCGAGTTTAGGTTTTTTCATTGCTTCTGCTGCTGCTGCTGCTGCTGCT 756
DB	704 ATCAAGTTTACCGAGTTTAGGTTTTTTCATTGCTTCTGCTGCTGCTGCTGCTGCTGCT 763
QY	757 AGCAGGCACCATGCTTCTGACGGCGAGCTGTACTTTTACAAATAGCTGCTGCTGCTGCTG 816
DB	764 AGCAGGCACCATGCTTCTGACGGCGAGCTGTACTTTTACAAATAGCTGCTGCTGCTGCTG 823
QY	817 CCGTCGTGTTGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
DB	824 CCGTCGTGTTGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883
QY	877 AGGATTTGCTATTCAATTTTCAAAAGAAATCTGAAAAAGAGAGCAACAACAAAGAACTGA 936
DB	884 AGGATTTGCTATTCAATTTTCAAAAGAAATCTGAAAAAGAGAGCAACAACAAAGAACTGA 943



Db 646 TGGGACTTTGACTTGGCTGCTACCTGGAAGAGCCTGAAGCCCGCAAGTCTGCAACTGT 705  
QY 634 AAATCTCACTGTGATTCGGTGTCCCAAGACACTGGAGGTGTATTAATATTCAGGTGT 693  
Db 706 AAATCTCACTGTGATTCGGTGTCCCAAGACACTGGAGGTGTATTAATATTCAGGTGT 765  
QY 694 ATTATCAAGTTTACCGAGTTAGGTTTTCATTCGCTACTTGGGCAAGTTGGACTTGG 753  
Db 766 ATTATCAAGTTTACCGAGTTAGGTTTTCATTCGCTACTTGGGCAAGTTGGACTTGG 825  
QY 754 ACTAGCAGCACCCTCTTCTGACGCCGACGTGTACTCTTACAATACGCTGCTGCTGTG 813  
Db 826 ACTAGCAGCACCCTCTTCTGACGCCGACGTGTACTCTTACAATACGCTGCTGCTGTG 885  
QY 814 CGCCGCTGCTGTTGTGGTGTCAACTGCTGCTGCGGTGTTGTTCTGCTGTAGAGAAA 873  
Db 886 CGCCGCTGCTGTTGTGGTGTCAACTGCTGCTGCGGTGTTGTTCTGCTGTAGAGAAA 945  
QY 874 AAGAGGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAAC 933  
Db 946 AAGAGGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAAC 1005  
QY 934 TGAGACAGAAAGTGAAATGAAATCCCGGTCTACAAATTCAGATGAACAAAGAACACACAGA 993  
Db 1006 TGAGACAGAAAGTGAAATGAAATCCCGGTCTACAAATTCAGATGAACAAAGAACACACAGA 1065  
QY 994 CACCGCTTCTCTCCCTCCCAATCTGTGAATCCAGTGAATCTGAAACAAAGAACACAGTAG 1053  
Db 1066 AACCGCTTCTCTCCCTCCCAATCTGTGAATCCAGTGAATCTGAAACAAAGAACACAGTAG 1125  
QY 1054 CTGTGGCCCTCTCTACCGCGGCTGATCAACGTCACCCAGCCGACGAAGTCATCCACA 1113  
Db 1126 CTGTGGCCCTCTCTACCGCGGCTGATCAACGTCACCCAGCCGACGAAGTCATCCACA 1185  
QY 1114 GCGTCTTTTAAATGCGCAGCTCTGAGAGTCAAGTCAATCAACTGTAGTATAG 1168  
Db 1186 GCGTCTTTTAAATGCGCAGCTCTGAGAGTCAAGTCAATCAACTGTAGTATAG 1240

RESULT 4  
ABK13031  
ID ABK13031 standard; cDNA; 1139 BP.  
AC ABK13031;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE DNA encoding human B7-like protein, B7-L\_h4.  
XX  
KW Human; B7-like protein; B7-L; antifertility; gynaecological;  
KW antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
KW antiasthmatic; nephrotoxic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 1..1134  
FT /\*tag= a  
FT /product= "B7-like protein, B7-L\_h4"

WO2000200710-A2.  
XX  
FN 03-JAN-2002.  
XX  
PD 28-JUN-2001; 2001WO-US20719.  
XX  
PF 28-JUN-2000; 2000US-214512P.  
XX  
PR 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.  
PA Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
XX WPI; 2002-130881/17.  
XX  
XX New B7-like polypeptides, polynucleotides and their modulators, useful  
PT for diagnosing, preventing and treating reproductive, immune and  
PT proliferative disorders, e.g. cancer and arteriosclerosis -  
XX Claim 1; Fig 4; 135pp; English.  
XX  
CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
CC useful for treating B7-like polypeptide-related disease, disorders or  
CC conditions including reproductive disorders (e.g. infertility, and  
CC miscarriage, preterm labour and delivery and endometriosis) and  
CC proliferative disorders. Antibodies, soluble proteins comprising  
CC extracellular domains and other regulators of B7-L polypeptides are  
CC useful for enhancing the immune response to tumours. (I) plays a role in  
CC growth and maintenance of cancer cells based on the observation of  
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
CC polypeptide. Hence modulators of (I) are useful for the treatment of  
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
CC in allograft transplantation, graft versus host disease, T-cell  
CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
CC molecules are useful for alleviating the symptoms associated with  
CC diseases involving chronic immune cell dysfunction or to treat  
CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
CC purpura and psoriasis, chronic inflammatory disease such as  
CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
CC are also useful as immunosuppressive agents for bone marrow and organ  
CC transplantation or to prolong graft survival. B7-L molecules are also  
CC useful for diagnosis and treatment of diseases involving abnormal cell  
CC proliferation, including arteriosclerosis and vascular restenosis.  
CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
CC shock syndrome or allosensitisation due to blood transfusions, and for  
CC treatment of allergy, asthma and hypersensitivity reactions,  
CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
CC lymphoproliferative disorders such as multiple myeloma. The present  
CC sequence represents the coding sequence of human B7-L\_h4.  
XX  
SQ Sequence 1139 BP; 290 A; 300 C; 283 G; 266 T; 0 other;  
Query Match 94.4%; Score 1102.8; DB 24; Length 1139;  
Best Local Similarity 97.9%; Pred. No. 1.9e-297;  
Matches 1136; Conservative 0; Mismatches 2; Indels 22; Gaps 1;  
QY 8 ATGTGGCAGGAGCCATGGAAAAATAGACCCACCCCGTCTCTGGTCTGTAATGAAGTC 67  
Db 1 ATGGTGGCAGGAGCCATGGAAAAATAGACCCACCCCGTCTCTGGTCTGTAATGAAGTC 60  
QY 68 ATAGAGGCCCCCAAAATGCAAGAGTCTTGAAGGGCTCCAGGGCTTCAACTGCACC 127  
Db 61 ATAGAGGCCCCCAAAATGCAAGAGTCTTGAAGGGCTCCAGGGCTTCAACTGCACC 120  
QY 128 GTCTCCCGAGGGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGGTGTGTAAGCGTC 187  
Db 121 GTCTCCCGAGGGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGGTGTGTAAGCGTC 180  
QY 188 AGGCCCATGGAGCCCATCATCAATGACCGCTTTCACCTCTCAGAGGTACGACAGGGC 247  
Db 181 AGGCCCATGGAGCCCATCATCAATGACCGCTTTCACCTCTCAGAGGTACGACAGGGC 240  
QY 248 GGGAACTTCACCTCGGAGATGATCATCCAAATGTGGAGCCCACTGATTCGGGGAACATC 307

Db 241 GGGAAACCTCACCTCGAGATGATCCACAATGTGGAGCCAGGATTCGGGGAAACATC 300  
 QY 308 AGATCAGCCTCCAGAAAGAGTCCGCTGATGATCTGCTTACCTACCGTCCAAAGTTATG 367  
 Db 301 AGATCAGCCTCCAGAAAGAGTCCGCTGATGATCTGCTTACCTACCGTCCAAAGTTATG 360  
 QY 368 GGAGAGCTGTTCAATCCAGTGTAACTTAACTTGTAGTCGTGAGAAATGAACCTTTGTAAGTT 427  
 Db 361 GGAGAGCTGTTCAATCCAGTGTAACTTAACTTGTAGTCGTGAGAAATGAACCTTTGTAAGTT 420  
 QY 428 ACTTGTCTACCTCAGCTGACCTGGCTCCCGATATTTCTCGAGCTCGTCTCCG 487  
 Db 421 ACTTGTCTACCTCAGCTGACCTGGCTCCCGATATTTCTCGAGCTCGTCTCCG 480  
 QY 488 GTCAGCCATTCAAGCTATTATTTGTCGGAGCCAGCAGCTTCAAAGTGCAGTGAGC 547  
 Db 481 GTCAGCCATTCAAGCTATTATTTGTCGGAGCCAGCAGCTTCAAAGTGCAGTGAGC 540  
 QY 548 ATCTGTGCTCTGACCCACAGACCAATGGACTTTGACTTGGCTGCTACCTGGAGAGC 607  
 Db 541 ATCTGTGCTCTGACCCACAGACCAATGGACTTTGACTTGGCTGCTACCTGGAGAGC 600  
 QY 608 CTGAAGGCCCGCAAGCTCTGCAACTGTAAATCTCACTGATTCGGTGTCCCAAGACACT 667  
 Db 601 CTGAAGGCCCGCAAGCTCTGCAACTGTAAATCTCACTGATTCGGTGTCCCAAGACACT 660  
 QY 668 GGAGTGTGTTAATATTCAGGTGTATATCAAGTTTACCGAGTTTAGTTTTTCATG 727  
 Db 661 GGAGTGTGTTAATATTCAGGTGTATATCAAGTTTACCGAGTTTAGTTTTTCATG 720  
 QY 728 CTTACTTGGGCAAGTTGGACTTGGACTAGCAGCAGCAGCTCTTCTGACGCGCAGCTG 787  
 Db 721 CTTACTTGGGCAAGTTGGACTTGGACTAGCAGCAGCAGCTCTTCTGACGCGCAGCTG 780  
 QY 788 ACTTTACAATACGCTGTGCTGCTGCCCGCTGCTGTGTGTGCTGCAACTGCTGCTGC 847  
 Db 781 ACTTTACAATACGCTGTGCTGCTGCCCGCTGCTGTGTGTGCTGCAACTGCTGCTGC 840  
 QY 848 CGTTGTTGTTCTGCTGTAGAGAAAGAGAGATTCGATTCATATTCAAAGAAATCT 907  
 Db 841 CGTTGTTGTTCTGCTGTAGAGAAAGAGAGATTCGATTCATATTCAAAGAAATCT 878  
 QY 908 GAAAAGAGAGACACAAAGAAAGTGAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 967  
 Db 879 GAAAAGAGAGACACAAAGAAAGTGAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 938  
 QY 968 AATTCAGATGAACAAAGAGACACAGACCGCTTCTCTCCCTCCCAATCTCTGTAATCC 1027  
 Db 939 AATTCAGATGAACAAAGAGACACAGACCGCTTCTCTCCCTCCCAATCTCTGTAATCC 998  
 QY 1028 AGTGATCTGAAAGAAAGAGAGTGTGCGCTCTCACCAGCGGCTGATCAAGT 1087  
 Db 999 AGTGATCTGAAAGAAAGAGAGTGTGCGCTCTCACCAGCGGCTGATCAAGT 1058  
 QY 1088 CCACCCAGCGCAGCAAGTCAATCCAGAGCTTCTTTTAATCTGCGCAGCTCTGAGAGGTC 1147  
 Db 1059 CCACCCAGCGCAGCAAGTCAATCCAGAGCTTCTTTTAATCTGCGCAGCTCTGAGAGGTC 1119  
 QY 1148 AGTAATACAACTGTAGTATA 1167  
 Db 1119 AGTAATACAACTGTAGTATA 1138

RESULT 5  
 AAS92356  
 ID AAS92356 standard; cDNA; 1392 BP.  
 XX  
 AC AAS92356;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #28160.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 XX P-PSDB; ABG28169.  
 DR New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX Claim 1; SEQ ID No 28160; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 1392 BP; 338 A; 361 C; 341 G; 352 T; 0 other;  
 Query Match 66.6%; Score 777.6; DB 23; Length 1392;  
 Best Local Similarity 87.2%; Pred. No. 1.4e-206;  
 Matches 912; Conservative 0; Mismatches 24; Indels 110; Gaps 1;  
 QY 152 ATGTGGGCTCTCAGTGACATGCTGTGCTAAGCGTCAGGCCCATGAGCCCATCATCACC 211  
 Db 1 ATGTGGGCTCTCAGTGACATGCTGTGCTAAGCGTCAGGCCCATGAGCCCATCATCACC 60  
 QY 212 AATGACCGCTTCACTCTCAGAGGTACGACAGGGGGGAACTTCACTCGAGATGATC 271  
 Db 61 AATGACCGCTTCACTCTCAGAGGTACGACAGGGGGGAACTTCACTCGAGATGATC 120  
 QY 272 ATCCACAATGTGGAGCCAGTGATTCGGGGAACTCAGATCGAGCTCCAGACAGTGC 331  
 Db 121 ATCCACAATGTGGAGCCAGTGATTCGGGGAACTCAGATCGAGCTCCAGACAGTGC 180  
 QY 332 CTGATGAGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGT 391  
 Db 181 CTGATGAGATCTGCTTACCTTACCGTCCAAAGTTATGGAGAGCTGTTCAATCCAGTGT 240



Matches		701;	Conservative	0;	Mismatches	356;	Indels	69;	Gaps	4;
Qy	42	CCGGTTCGGGTC	CTGGTAATGAAGTCATAGAGGCCCCAAAATGCAAGAGTCCTGAAGG	101						
Db	108	CAGCTTCGGATCCAGTTATCAGATCATAGAAGTCTCAGAAATGTAACAGTCTCTAAGG	167							
Qy	102	GCTCCAGGCTCGCTTCAACTGCAACCGTCTCCAGGCGTGAAGCTCATCTGTGGGTC	161							
Db	168	ACTCAGAGGCTCACCTTCAACTGCACCGTGACTCAGCGGTGGAAGCTTCTCATGTGGACTC	227							
Qy	162	TCAGTGACATGGTGTCTAAGCTCAGGCCATGAGCCCATCATCATCAACATGACCGCT	221							
Db	228	TTAACCAAAATGGTGTGCTGAGTCTCACCAACCAAGGACCCATCATCACCAACACCGCT	287							
Qy	222	TCACCTCTCAGAGGTACACCAAGCGGGGAATCTCACCTCGGAGATGATCATCCACAATG	281							
Db	288	TCACCTATGCCAGITTACAAAGCACTGACAGCTTCATCTCGAGTTGATCATCATGATG	347							
Qy	282	TGGAGCCAGTGATTCGGGGAAACATCAGATGAGCCTCCAGAACAGTGCCTTGAATGGAT	341							
Db	348	TGCAGCCAGTGACTCGGGATCCCGTCAATGCAGCTGCAGAACAGCCATGGGTTGGAT	407							
Qy	342	CTGCTTACTCTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAACTTTGTAG	401							
Db	408	CTGCCCTCTCAGTGAAGTATGAGGACCCCTGAACATTTCTTAGCAACACCTTATAG	467							
Qy	402	TCGCTGAGAATGAACCTTTGTGAAGTTACTTGTCTACCCCTCACACTGACCTGGCTGCCGG	461							
Db	468	TCACGTAGGGTGAACCTGTATGTGACTTGTATGCGGTGGCTGGACCTCACTCCCGG	527							
Qy	462	ATATTTCTGGAGCTCGTCTCTGCTGCTGAGCCATCAAGCTATTAATTTGTTCCGGAGC	521							
Db	528	ATATTTCTGGAGCTTGAGGTTCOCGTGAAGCCATTTCAGATTACAAATTCCTTTCTGGAGC	587							
Qy	522	CCAGCGACTTCAAAGTGCAGTGACATCTCTGGCTCTGACCCCAAGCAATGGGACTT	581							
Db	588	CGGGCACTTTATGAGGCTTTGAGTGTCTTGACCTCACACCCTGGGCAAGGGACT	647							
Qy	582	TGACTTGGCTGCTACCTGGAGAGCCTGAAGGCCCGCAAGTCTGCAACTGTAAATCTCA	641							
Db	648	TGACTTGTGTGGCAGAGCTGAAGGACTTGCAGGCCCAAGTCTTAACTGTCAACCTGA	707							
Qy	642	CTGTGATTCGGTGTCCCAAGACACTGGAGGTGTTAATAATTCAGGTGTATTATCAA	701							
Db	708	CTGTGTTTCAGCCTCCACTGACAGATTGGAGAGGA-----	744							
Qy	702	GTTTACCGAGTTAGTGTTCCTTCAATGCGCTACTTGGGGCAAGTTGGACTTGGACTAGCAG	761							
Db	745	-----AGGCCAGCACTGCCGACCTGGGCCATCATCTGCTGGCAGTGGCT	791							
Qy	762	GCACCATGCTTGTGACGGGAGCTGTACTTTCATAATAGCTGTGCTGTGCGCCCGTC	821							
Db	792	TTTCCTTGTCTGTGATCTGATCATTTGTTTGAATATAATATTCTGTGCTG-----	843							
Qy	822	GTTGTGTGGCTGCAACTGCTGCTCGCGTGTGTTGTTCTGTGTGAGAAAGAGGAT	881							
Db	844	-----TTGTGCTCCAGAGAGAAAGGAGAAATCTACTTATCAAAATGAATAGGAA	898							
Qy	882	TTGCTATTCAATTTCAAAGAAATCTGAAGAGAGAGACAAACAAAGAACTGTGACAG	941							
Db	899	TCTGCAAAACATGAGACAA-----ACAAAGCAGATCCGGAGACAAAGTTAA	944							
Qy	942	AAAGTGGAAATGAAGCTCCGCTCAATTCAGATGACAAAGACCAAGACCGCTT	1001							
Db	945	AAAGTGGAAAGAAACATACGGGTACAGTTTCGGATGAGGCAAGGCTGCACAGACTGCAT	1004							
Qy	1002	CTCTCCCTCCCAATCTCTGTAATCCAGTGATCTCTGAACAAAGAAACAGTAGCTGTGCC	1061							
Db	1005	CTCTCCCTCCCAATCTCTGTAAGTCAGCTTCCAGAGAAAGCGACAGTA-----GCC	1058							
Qy	1062	CTCTCAACAGGGGCTGATCAACCTCCACCCAGGCCAGAGTCAATCCACAGGTTCTT	1121							
Db	1059	TTCTTATCAGGAACCTCAATAAACATCATCCCGGTCAGCACTCATCCAGGGTTCTT	1118							

Qy	1122	TTAATCTGGCCAGTCTCTGAGAGGTCAAGTAATACAACACTGTAGTATA	1167							
Db	1119	TTGACATCGCCAGTCTCTCAGAGGTCAAGATGTGACTTTTAGTGTA	1164							
RESULT 7										
AAI36582										
ID	AAI36582	standard; DNA; 401 BP.								
XX	AAI36582;									
AC	AAI36582;									
XX	17-OCT-2001	(first entry)								
DT										
XX										
DE	Probe #5268	used to measure gene expression in human placenta sample.								
XX										
KW	Probe; microarray; human; placenta; antenatal diagnosis;									
KW	genetic disorder; ss.									
XX	Homo sapiens.									
OS	WO200157272-A2.									
PN										
XX	09-AUG-2001.									
PD										
XX	30-JAN-2001; 2001WO-US00663.									
PF										
XX	04-FEB-2000; 2000US-0180312.									
PR	26-MAY-2000; 2000US-0207456.									
PR	30-JUN-2000; 2000US-0608408.									
PR	03-AUG-2000; 2000US-0632366.									
PR	21-SEP-2000; 2000US-0234687.									
PR	27-SEP-2000; 2000US-0236359.									
PR	04-OCT-2000; 2000GB-0024263.									
XX	(MOLE-) MOLECULAR DYNAMICS INC.									
PA										
XX	Penn SG, Hanzel DK, Chen W, Rank DR;									
PI										
XX	WPI; 2001-488897/53.									
DR										
XX	Human genome-derived single exon nucleic acid probes useful for									
PT	analyzing gene expression in human placenta -									
XX										
PS	Claim 25; SEQ ID No 5268; 654pp; English.									
XX	The present invention relates to single exon nucleic acid probes (SENPs).									
CC	The present sequence is one such probe. The probes are useful for									
CC	producing a microarray for predicting, measuring and displaying gene									
CC	expression in samples derived from human placenta. The probes are useful									
CC	for antenatal diagnosis of human genetic disorders.									
XX										
SQ	Sequence 401 BP; 105 A; 98 C; 111 G; 87 T; 0 other;									
Query Match										
Best Local Similarity		16.8%;	Score	196.8;	DB	22;	Length	401;		
Matches		201;	Conservative	0;	Mismatches	7;	Indels	0;	Gaps	0;
Qy	27	AAAATAGAGACCCACCCCGTTCTCGGCTCTGCTGTAATGAAGTCATAGAAGGCCCCCAAAATG	86							
Db	194	AAATGGCTACTCTCCAGTTCTGGTCTGTAATGAAGTCATAGAAGGCCCCCAAAATG	253							
Qy	87	CAAGAGTCTGAAGGGTCCCGAGCTCGCTTCAACTGCACTGCTCCAGGGCTGGAAGC	146							
Db	254	CAAGAGTCTGAAGGGTCCCGAGCTCGCTTCAACTGCACTGCTCCAGGGCTGGAAGC	313							
Qy	147	TCATCATGTGGGCTCTCAGTCACATGCTGCTGAAGCTCAGGCCCATGAGCCCATCA	206							
Db	314	TCATCATGTGGGCTCTCAGTCACATGCTGCTGAAGCTCAGGCCCATGAGCCCATCA	373							
Qy	207	TCACCAATGACCGCTTCACTCTCTCAGAG	234							
Db	374	TCACCAATGACCGCTTCACTCTCTCAGAG	401							

## RESULT 8

ABK13034  
ID ABK13034 standard; cDNA; 754 BP.

AC ABK13034;

XX 23-APR-2002 (first entry)

XX DNA encoding mouse B7-like protein, B7-L\_m3.

XX Mouse; B7-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.

XX Mus musculus.

XX Key Location/Qualifiers  
XX CDS 53..724

XX FT /\*tag= a

XX FT /product= "B7-like protein, B7-L\_m3"

XX PN WO200200710-A2.

XX PD 03-JAN-2002.

XX PF 28-JUN-2001; 2001WO-US20719.

XX PR 28-JUN-2000; 2000US-214512P.

XX PR 28-NOV-2000; 2000US-0729264.

XX PA (AMGE-) AMGEN INC.

XX PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX WPI; 2002-130881/17.

XX DR P-PSDB; AAU75546.

XX New B7-like polypeptides, polynucleotides and their modulators, useful

XX for diagnosing, preventing and treating reproductive, immune and

XX proliferative disorders, e.g. cancer and arteriosclerosis

XX Claim 1; Fig 7; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
XX useful for treating B7-like polypeptide-related disease, disorders or  
XX conditions including reproductive disorders (e.g. infertility,  
XX miscarriage, preterm labour and delivery and endometriosis) and  
XX proliferative disorders. Antibodies, soluble proteins comprising  
XX extracellular domains and other regulators of B7-L polypeptides are  
XX useful for enhancing the immune response to tumours. (I) plays a role in  
XX growth and maintenance of cancer cells based on the observation of  
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
XX polypeptide. Hence modulators of (I) are useful for the treatment of  
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
XX in allograft transplantation, graft versus host disease, T-cell  
XX dependent B-cell mediated diseases and autoimmune diseases. B7-L  
XX molecules are useful for alleviating the symptoms associated with  
XX diseases involving chronic immune cell dysfunction or to treat  
XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
XX purpura and psoriasis, chronic inflammatory disease such as  
XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They

CC are also useful as immunosuppressive agents for bone marrow and organ  
CC transplantation or to prolong graft survival. B7-L molecules are also  
CC useful for diagnosis and treatment of diseases involving abnormal cell  
CC proliferation, including arteriosclerosis and vascular restenosis.  
CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
CC shock syndrome or allosensitisation due to blood transfusions, and for  
CC treatment of allergy, asthma and hypersensitivity reactions,  
CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
CC lymphoproliferative disorders such as multiple myeloma. The present  
CC sequence represents the coding sequence of mouse B7-L\_m3.

XX Sequence 754 BP; 220 A; 191 C; 175 G; 168 T; 0 other;

XX Query Match 15.6%; Score 181.8; DB 24; Length 754;  
XX Best Local Similarity 72.9%; Pred. No. 2.7e-40;  
XX Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 42 CCGGTTCTGGGTCTGGTAATGAAGTCATAGAAGGCCCCCAAAATGCAAGAGTCTCTGAAGG 101  
DB 108 CAGCTTCGGGATCCAGTTATCAGATCATAGAAGGTCTCAGAATGTAACAGTCTCTAAAGG 167  
QY 102 GCTCCAGGCTCGTTCAACTGCACGGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTC 161  
DB 168 ACTCAGAGGCTCACTTCAACTGACCGTGACTCACGGCTGGAAGCTTCTCATGTGGACTC 227  
QY 162 TCAGTGACATGTTGGTGTAAAGCTCAGGCCCATGAGGCCCATCATCAACCAATGACCGCT 221  
DB 228 TTAACCAATGTTGGTGTGAGTCTCACCACCAAGGACCATCATCAACCAACCGCT 287  
QY 222 TCACCTCTCAGAGGTACGACCGGGCGGAACCTTCACTCGGAGATGATCATCCACAATG 281  
DB 288 TCACCTATGCCAGTTACAAACAGCAGCTGACAGCTTCACTCTCGAGTTGATCATCATGATG 347  
QY 282 TGGAGCCAGTGATTGGGGGAACATCAGATGAGCGCTCCAGAACAGTCTCGCTCATGGAT 341  
DB 348 TGCAGCCAGTGACTCGGGATCCGTGCAATGAGCGCTGCAAGACAGCCATGGGTTTGGAT 407  
QY 342 CTGCTTACCTTACCGTCCAAAG 362  
DB 408 CTGCTTCTCTCAGTGCAAG 428

## RESULT 9

ABK13033

ID ABK13033 standard; cDNA; 895 BP.

XX AC ABK13033;

XX DT 23-APR-2002 (first entry)

XX DE DNA encoding mouse B7-like protein, B7-L\_m2.

XX Mouse; B7-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder; gene; ss.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

XX CDS 53..865

XX FT /\*tag= a

XX FT /product= "B7-like protein, B7-L\_m2"

XX PN WO200200710-A2.

PD	03-JAN-2002.		
XX			
PF	28-JUN-2001; 2001WO-US20719.	Db	228 TTAACAAATGGTGTGTGTGAGTCTCAACCAAGGACCCATCATCAACAACACCGCT 287
XX		QY	222 TCACCTCTCAGAGGTACGACACGAGGCGGGAACCTTCACCTCGGAGATGATCATCCCAATG 281
XX	28-JUN-2000; 2000US-214512P.	Db	288 TCACCTATGCCAGTTACAAACAGCACTGACAGCTTCATCTCGGAGTTGATCATCCATG 347
PR	28-NOV-2000; 2000US-0729264.	QY	282 TGGAGCCCGAGTATTGGGGGAACATCAGATGAGGCTCCAGAACAGTGCCCTGCGATGGAT 341
XX	(AMGE-) AMGEN INC.	Db	348 TCGAGCCCGAGTACCTCGGGATCCGTGCAATGAGCTGCAGAACAGCCATGGGTTGGAT 407
PA		QY	342 CTGCTTACCTTACCGTCCAAAG 362
PI	Welcher AA, Sarmiento UM, Schultz HU, Chute HT;	Db	408 CTGCTTCTCTCAGTCAAG 428
XX			
DR	WPI: 2002-130881/17.		
DR	P-PSDB; AAU75345.		
XX			
PT	New B7-like polypeptides, polynucleotides and their modulators, useful		
PT	for diagnosing, preventing and treating reproductive, immune and		
PT	proliferative disorders, e.g. cancer and arteriosclerosis -		
XX			
PS	Claim 1; Fig 6; 135pp; English.		
XX			
CC	The invention relates to an isolated B7-like (B7-L) polypeptide (I).		
CC	The polypeptide, polynucleotide encoding it and antibody against (I) are		
CC	useful for treating B7-like polypeptide-related disease, disorders or		
CC	conditions including reproductive disorders (e.g. infertility,		
CC	miscarriage, preterm labour and delivery and endometriosis) and		
CC	proliferative disorders. Antibodies, soluble proteins comprising		
CC	extracellular domains and other regulators of B7-L polypeptides are		
CC	useful for enhancing the immune response to tumours. (I) plays a role in		
CC	growth and maintenance of cancer cells based on the observation of		
CC	seminal vesicle hyperplasia in transgenic mice overexpressing B7-L		
CC	polypeptide. Hence modulators of (I) are useful for the treatment of		
CC	cancer including seminal vesicle cancer, lung, brain, breast, ovarian,		
CC	testicular cancer and cancers of haematopoietic system. B7-L polypeptide		
CC	pathway can be manipulated to regulate cytotoxic T-lymphocyte response		
CC	in allograft transplantation, graft versus host disease, T-cell		
CC	dependent B-cell mediated diseases and autoimmune diseases. B7-L		
CC	molecules are useful for alleviating the symptoms associated with		
CC	diseases involving chronic immune cell dysfunction or to treat		
CC	autoimmune diseases such as systemic lupus erythematosus, rheumatoid		
CC	arthritis, multiple sclerosis, diabetes, immune thrombocytopenic		
CC	purpura and psoriasis, chronic inflammatory disease such as		
CC	inflammatory bowel disease (Crohn's disease and ulcerative colitis),		
CC	Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They		
CC	are also useful as immunosuppressive agents for bone marrow and organ		
CC	transplantation or to prolong graft survival. B7-L molecules are also		
CC	useful for diagnosis and treatment of diseases involving abnormal cell		
CC	proliferation, including arteriosclerosis and vascular restenosis.		
CC	Antagonists of B7-L polypeptides are useful for alleviation of toxic		
CC	shock syndrome or allosensitisation due to blood transfusions, and for		
CC	treatment of allergy, asthma and hypersensitivity reactions,		
CC	nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and		
CC	extrinsic alveolitis), vasculopathies (Grave's disease), various pneumopathies		
CC	(thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and		
CC	lymphoproliferative disorders such as multiple myeloma. The present		
CC	sequence represents the coding sequence of mouse B7-L <sub>m2</sub> .		
XX			
SQ	Sequence 895 BP; 250 A; 223 C; 212 G; 210 T; 0 other;		
	Query Match 15.6%; Score 181.8; DB 24; Length 895;		
	Best Local Similarity 72.9%; Pred. No. 3e-40;		
	Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;		
QY	42 CCGGTTCTGGGCTGTATGAGTCATAGAGGCCCCAAATCCAGAGTCCTCAAGG 101		
Db	108 CAGCTTCGGATCCAGTTATCAGATCATAGAGGTCCTCAGATGTACAGTCCTAAAGG 167		
QY	102 GCTCCAGGCTCGCTTCAACTCAGCGTCTCCAGGGCTGGAAGCTCATCATGTGGGTC 161		
Db	168 ACTCAGAGGCTCAGTCACTCAGCGTACATCAGGCTGGAGCTTCTCATGTGACTC 227		
QY	162 TCAGTGACATGTGTGCTAAGCGTCAGGCCCATGAGCCCATCATCAACATGACCGCT 221		

RESC<sup>h</sup>LT 10  
ABX55016  
ID ABX55016 standard; cDNA; 398 BP.  
XX  
AC ABX55016;  
XX  
DT 26-FEB-2003 (first entry)  
XX  
DE Bovine EST associated with lactation/muscle/fat deposition #4945.  
XX  
KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
OS Bos Taurus.  
XX  
PN US2002137160-A1.  
XX  
PD 26-SEP-2002.  
XX  
PF 26-OCT-2001; 2001US-0983965.  
XX  
PR 17-DEC-1998; 98US-113678P.  
PR 15-DEC-1999; 99US-0465231.  
XX  
PA (BYAT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX  
PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
WPI; 2003-102386/09.  
XX  
PT Purified nucleic acid molecules, useful for genome mapping, gene  
PT identification and analysis, cattle breeding or preparation of  
PT constructs for cattle gene expression and genetically improved cattle -  
PS Claim 2; SEQ ID No 4945; 38pp; English.  
XX  
CC The invention relates to a purified nucleic acid molecule associated with  
CC lactation or muscle and fat deposition (designated LMPD), derived  
CC from cattle, and the LMPD nucleic acid can specifically hybridise to a  
CC second nucleic acid molecule comprising any of 5912 nucleotide  
CC sequences, appearing as ABX50072-ABX55983, or complements of them.  
CC Also included are: (1) a transformed cell having a nucleic acid  
CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-  
CC translated sequence that functions in the cell to cause termination of  
CC transcription and addition of polyadenylated ribonucleotides to a 3' end  
CC of the mRNA molecule; and (2) determining a level or pattern of a  
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker  
CC nucleic acid (comprising any of the 5912 nucleic acid sequences or its  
CC complement or fragment) with a complementary nucleic acid molecule  
CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 5912 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137160.  
XX  
SQ Sequence 398 BP; 88 A; 102 C; 137 G; 71 T; 0 other;  
Query Match 11.7%; Score 136.2; DB 25; Length 398;  
Best Local Similarity 77.5%; Pred. No. 1e-27;  
Matches 165; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 42 CCGGTTCTGGTCTGGTAAATGAAGTCATAGAGGCCGCCCAAAATGCAAGAGTCCTGAAGG 101  
DB |||||  
DB 186 CAGCCTGTGGATCCAGCAGTGAATCATAGAGGGTCCCAAGATGTCACAGCCCTGAAGG 245  
QY 102 GCTCCAGGCTCGCTTCAACTGACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTC 161  
DB |||||  
DB 246 GCTCGAGGCTCGCTTCAACTGACCATCTCGAGGGCTGGAAGCTCGTCATGTGGGCTC 305  
QY 162 TCAGTGAATGTGTGTGCTAAGCGTCAGGCCCATGGAGCCCATCATCAACCAATGACCGCT 221  
DB |||||  
DB 306 TCAGAGGCACATGTGTGCTGACATGACACCTTAATGAGACCATCATCACCAGTGACCGCT 365  
QY 222 TCACCTCTAGAGGTACGACGAGCGGGGAACT 254  
DB |||||  
DB 366 TCACCTTCGCAAGCTACCAAGAGGGCGGAACT 398  
RESULT 11  
ABK10239/c  
ID ABK10239 standard; DNA; 277 BP.  
XX  
AC ABK10239;  
XX  
DT 20-MAY-2002 (first entry)  
XX  
DE Trinucleotide repeat sequence #2.  
XX  
KW Trinucleotide repeat; fragile X syndrome; ds; DRPLA;  
KW spinocerebellar ataxia type III; Marfan syndrome;  
KW hereditary hypertrophic cardiomyopathy; neuropsychiatric;  
KW dentatorubral and pallidolysian atrophy;  
KW spinocerebellar ataxia; X-linked spinobulbar atrophy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT repeat\_region 152..194  
FT /\*tag= a  
FT /\*rpt\_type= TANDEM  
FT repeat\_unit 152..154  
FT /\*tag= b  
FT /\*note= "CAG type repeat"  
FT repeat\_region 198..222  
FT /\*tag= a  
FT /\*rpt\_type= TANDEM  
FT repeat\_unit 198..200  
FT /\*tag= b  
FT /\*note= "CAG type repeat"  
XX  
OS Homo sapiens.  
XX KR2000003004-A.  
XX  
PD 15-JAN-2000.  
XX  
PF 25-JUN-1998; 98KR-0024064.  
XX

PR 25-JUN-1998; 98KR-0024064.  
XX (SMSU ) SAMSUNG MEDICAL CENT.  
PA (JIND/) JIN D G.  
XX  
PI Jin DG;  
XX  
DR WPI; 2000-862424/64.  
XX  
PT Plasmid vector arraying neuropsychiatric gene more than copy number  
PT containing trinucleotide repeats its use thereof -  
XX  
PS Disclosure; Page 12; 23pp; Korean.  
XX  
CC This invention relates to the use of a plasmid vector array  
CC containing a neuropsychiatric gene containing trinucleotide repeats.  
CC Trinucleotide repeats and their expansion are known to be involved  
CC in a number of diseases including fragile X syndrome, spinocerebellar  
CC ataxia type III, hereditary hypertrophic, cardiomyopathy, dentatorubral  
CC and pallidolysian atrophy (DRPLA), spinocerebellar ataxia, Marfan  
CC syndrome, X-linked spinobulbar atrophy. The present sequence  
CC represents a nucleotide sequence of the invention showing the  
CC the CAG tri-nucleotide repeats.  
XX  
SQ Sequence 277 BP; 80 A; 66 C; 62 G; 69 T; 0 other;  
Query Match 4.6%; Score 53.4; DB 21; Length 277;  
Best Local Similarity 59.6%; Pred. No. 0.00012;  
Matches 90; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
QY 801 GCTGCTGCTGCTGCCGCGTCTGTTGTTGGCTGCAACTGCTGCTGCCCTTGTGTTCT 860  
DB |||||  
DB 215 GCTGCTGCTGCTGCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156  
QY 861 GCTGTAGAAGAAAAAGAGATTTCGTATTCAATTTCAAAAGAAATCTGAAAAAGAGAAGA 920  
DB |||||  
DB 155 GCTGTTGCTGCTTTTGTGCTGCTCTGAAACATTCAAAAGTGAAGTATATTTAAAAACAA 96  
QY 921 CAACAAGAAACTGAGACAGAAAGTGGAAA 951  
DB |||||  
DB 95 AACTTAAAGAAATAAATACACCATGAGAAA 65  
RESULT 12  
ABL29756  
ID ABL29756 standard; DNA; 7029 BP.  
XX  
AC ABL29756;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40741.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 23:23:44 ; Search time 9.26863 Seconds  
(without alignments)  
1938.172 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059

Sequence: 1 MGLVIFLHSGSGNEIVEGP.....HPQASFNLASPEKVSNTTVV 382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	147.5	7.2	404	1	RAGE HUMAN
2	131	6.4	337	1	G55A_CHICK
3	129.5	6.3	344	1	NTRI_RAT
4	128	6.2	353	1	CEPU_CHICK
5	128	6.2	1447	1	DCC_HUMAN
6	128	6.2	1914	1	KMLS_HUMAN
7	127	6.2	1447	1	DCC_MOUSE
8	125.5	6.1	6632	1	UN89 CAEL
9	124	6.0	1070	1	PTK7_HUMAN
10	123.5	6.0	344	1	NTRI_HUMAN
11	123.5	6.0	344	1	NTRI_MOUSE
12	120	5.8	345	1	OPCM_HUMAN
13	119	5.8	345	1	OPCM_BOVIN
14	118	5.7	345	1	OPCM_RAT
15	117.5	5.7	1443	1	NEO1_CHICK
16	116	5.6	4391	1	PGSM_HUMAN
17	114.5	5.6	1461	1	NEO1_HUMAN
18	114	5.6	569	1	SILF_MOUSE
19	113.5	5.5	416	1	RAGE_BOVIN
20	113.5	5.5	1377	1	NEO1_RAT
21	112	5.4	764	1	ICCR_DROME
22	110.5	5.4	524	1	BUTY_MOUSE
23	110.5	5.4	620	1	SMP_COTJA
24	109.5	5.3	249	1	CSP_DROME
25	107	5.2	365	1	PGSM_MOUSE
26	107	5.2	3707	1	PGSM_MOUSE
27	106.5	5.2	333	1	AMAL_DROME
28	106	5.1	319	1	A33_HUMAN
29	106	5.1	1092	1	NCA2_XENLA
30	104.5	5.1	862	1	CD22_MOUSE
31	104.5	5.1	879	1	FPRP_RAT
32	104	5.1	879	1	FPRP_MOUSE
33	103.5	5.0	338	1	LAMP_CHICK

## RESULT 1

ID	RAGE HUMAN	STANDARD;	PRT;	404 AA.
AC	Q15109; Q15279; Q9H2X7; Q9Y3R3;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Advanced glycosylation end product-specific receptor precursor			
DE	(Receptor for advanced glycosylation end products).			
GN	AGER OR RAGE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Lung;			
RX	MEDLINE=92340547; PubMed=1378943;			
RA	Nepper M. Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,			
RA	Elliston K., Stern D., Shaw A.;			
RT	"Cloning and expression of a cell surface receptor for advanced			
RT	glycosylation end products of proteins.";			
RL	J. Biol. Chem. 267:14998-15004(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=95137587; PubMed=7835890;			
RA	Sugaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,			
RA	Inoko H., Ikemura T.;			
RT	"Three genes in the human MHC class III region near the junction with			
RT	the class II: gene for receptor of advanced glycosylation end			
RT	products, PBX2 homeobox gene and a notch homolog, human counterpart			
RT	of mouse mammary tumor gene int-3.";			
RL	Genomics 23:408-419(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RA	Rowen L., Danks C., Baskin D., Faust J., Loretz C., Ahearn M.E.,			
RA	Banta A., Spies T., Hood L.;			
RT	"Sequence determination of 300 kilobases of the human class III MHC			
RT	locus.";			
RL	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.			
RA	Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,			
RA	Yamamoto H.;			
RT	"Molecular heterogeneity of the receptor for advanced glycation			
RT	endproducts.";			
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RA	Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,			
RA	Schuer A., Huber G.;			
RT	"cDNA cloning of a novel secreted isoform of the human receptor for			
RT	advanced glycation end products (AGEs) and characterization of cells			
RT	co-expressing cell-surface scavenger receptors and Swedish mutant			
RT	amyloid precursor protein.";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			

## ALIGNMENTS

34	103.5	5.0	890	1	TYO3_HUMAN	Q06418 homo sapien
35	102.5	5.0	344	1	CD2_RAT	P08921 rattus norv
36	102.5	5.0	348	1	KILO_RAT	Q920J8 rattus norv
37	102.5	5.0	365	1	CXAR_MOUSE	P97792 mus musculu
38	102.5	5.0	2499	1	MPRI_BOVIN	P08169 bos taurus
39	102	5.0	880	1	TYO3_RAT	P55146 rattus norv
40	100.5	4.9	442	1	SILF_HUMAN	O43699 homo sapien
41	100.5	4.9	873	1	FA52_DROME	P34082 drosophila
42	100	4.9	1241	1	NPHN_HUMAN	O60500 homo sapien
43	99.5	4.8	261	1	KLK8_RAT	P36374 rattus norv
44	99.5	4.8	1260	1	CAML_MOUSE	P11627 mus musculu
45	99.5	4.8	1906	1	KMLS_CHICK	P11799 gallus gall

```

RN RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[7]
RN RP SEQUENCE OF 1-12 FROM N.A.
RA Hudson B.I., Futers T.S.;
RT "Novel polymorphisms in the receptor for advanced glycation
RT end-products (RAGE) gene";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates interactions of advanced glycosylation end
CC products (AGE). These are nonenzymatically glycosylated proteins
CC which accumulate in vascular tissue in aging and at an accelerated
CC rate in diabetes. Receptor for amyloid beta peptide.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).
CC -!- SECRETED (isoform 2).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q15109-1; SequencesDisplayed;
CC Name=2; Synonyms=RAGESEC;
CC IsoId=Q15109-2; Sequences=VSP_002551, VSP_002552;
CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21211; AAA03574.1; -
CC EMBL; D28769; BAA05958.1; -
CC EMBL; U89336; AAB47491.1; -
CC EMBL; AB036432; BAA89369.1; -
CC EMBL; AJ133822; CAB43108.1; -
CC EMBL; BC020669; AAB20669.1; -
CC EMBL; AF208289; AAG35728.1; -
CC EMBL; I61596; I61596.
CC Genew; HGNC:320; AGER.
CC MIM; 600214; -
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004888; F:transmembrane receptor activity; TAS.
CC GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; TAS.
CC GO; GO:0006954; P:inflammatory response; TAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC Pfam; PF00047; Ig_2.
CC SMART; SM00408; Igc2; 1.
DR PROSITE; PS0835; IG LIKE; 3.
DR PROSITE; PS00290; IG MHC; 1.
DR Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
KW Alternative splicing; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 404
FT DOMAIN 23 342
FT TRANSMEM 343 363
FT DOMAIN 364 404
FT DOMAIN 23 116
FT DOMAIN 124 221
FT DOMAIN 227 317
FT DISULFID 38 99
FT DISULFID 144 208
FT DISULFID 259 301
FT CARBOHYD 25 25
FT CARBOHYD 81 81
FT DOMAIN 380 384
FT VARSPLIC 54 67
FT VARSPLIC 275 404
FT VARIANT 100 100
FT CONFLICT 1 1
FT SEQUENCE 404 AA; 42802 MW; 0D584C436C30CE7 CRC64;
SQ
Query Match 7.2%; Score 147.5; DB 1; Length 404;
Best Local Similarity 22.6%; Pred. No. 0.00018;
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;
QY 15 EVIEGPQNAVILKGSQARENCTVSQ---WKLIIMNALSMMVLSVSRPMPEITNDRFTS- 70
Db 125 EIVDSASELTA--GVPNKVGTCVSEGSYPAGTLSWHLDG-----KPLVPNEKGVSV 173
QY 71 ----QRYDOGNFT--SEMIHNVPSDSGNR-----CSLQNSRLHGSAYLTQVMGELF 120
Db 174 KEQTRHRHPETGLTQLQSELM--VTPARGDPRPTFCFSFSPCLPRHRLRTAPIQPRVW 230
QY 121 IP-----SVNLVVAENP-----CEVTCPLSHWTRLPDISWELGLLVSHSSY 162
Db 231 EPVPLEEVQLV--EPEGGAVAPGTVLTCEVPAQS-----PQIHWKMD----- 274
QY 163 YFVPEPSDLQSAVSITALTPOSNGLTCTVATWKSARKSATVNLTVIRCPQDTGGGINI 222
Db 275 -GVPLPLPPSPVLLIPGIPQDQGYSCVATHSHSGPQESRAVSIIE-PGEEG----- 327
QY 223 PGVLSLPSLGSFLPTWGVGLGLAGTMLLT-----PTCTLTIRCCCRRCGCCGCC 275
Db 328 -----PTAGSVGSGGLGTALALGILGGLGTAAALLIGVILWQR- 366
QY 276 CRCCFCRRKRGFRIOFKKSEKKT--NKETETESNGNSGYNSEDEQKTTDTASLPKKS 333
Db 367 -----QRRGERKAPQEEERAEIN----- 389
QY 334 CESSDPEQRNSCGPP 349
Db 390 -QSEEPAGESTGGP 404
RESULT 2
GSSA_CHICK STANDARD; PRT; 337 AA.
ID G55A_CHICK
AC Q98892;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

DE Neurite inhibitor GP55-A precursor (OBCAM protein gamma isoform).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Lodge A.P., Kim D.-S., Howard M.R., McNamee C.J., Smith N., Moss D.J.;  
RT "Cloning of CEPU-s, a secreted isoform of CEPU-1, and OBCAM cDNAs from  
RT chick: structural diversity of IGLON family proteins.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 64-337 FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Brain;  
RA MEDLINE=97157768; PubMed=9004047;  
RT "A family of glycoproteins (GP55), which inhibit neurite outgrowth,  
RT are members of the Ig superfamily and are related to OBCAM,  
RT neurotrophin, LAMP and CEPU-1.";  
RL J. Cell Sci. 109:3129-3138(1996).  
CC -!- FUNCTION: INHIBITS NEURITE OUTGROWTH.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO THE NERVOUS SYSTEM.  
CC -!- DEVELOPMENTAL STAGE: INCREASES DURING DEVELOPMENT FROM VERY LOW  
CC LEVELS AT EMBRYONIC DAY 10 AND IS MOST ABUNDANT AFTER HATCHING.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Y08170; CAB41420.1; -.  
CC InterPro; IPR007110; IG-like.  
CC InterPro; IPR003598; IG\_C2.  
CC InterPro; IPR003006; IG\_MHC.  
CC Pfam; PF00047; Ig; 3.  
CC SMART; SM00408; IGC2; 2.  
CC PROSITE; PS00835; IG LIKE; 3.  
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
CC Repeat; Signal.  
CC BY SIMILARITY.  
CC SIGNAL 1 20  
CC CHAIN 21 317  
CC PROPEP 318 337  
CC DOMAIN 32 119  
CC DOMAIN 129 211  
CC DOMAIN 215 302  
CC DISULFID 50 108  
CC DISULFID 150 194  
CC DISULFID 236 288  
CC CARBOHYD 133 133  
CC CARBOHYD 277 277  
CC CARBOHYD 285 285  
CC CARBOHYD 298 298  
CC SEQUENCE 337 AA; 36887 MW; BAE717551856651E CRC64;  
Query Match 6.4%; Score 131; DB 1; Length 337;  
Best Local Similarity 25.8%; Pred. No. 0.0029;  
Matches 65; Conservative 33; Mismatches 96; Indels 58; Gaps 14;  
QY 22 NATVLKGSQARFNTVSGWGLIMWLSDMVL-----SVRPMEPIITNDRTSQRYD 74  
Db 37 NVTVQGESATRLCTVDDRVRRAW-LNRSTILVAGNDKWSIDNRVLSN---TKTQY- 91  
QY 75 QGGNFTSEMIHNVPEPSDGNIRCSLQ-----NSRLHGSAYLTAVQVMGEIFIPSVNLV 128  
Db 92 -----SIKIHNVVDYDEGPTCSVQTDNHPKTRVH-----LIVQVPPQIWNISSDITV 140  
QY 129 AENEPCEVTCL-----PSHWRLPDISWELGLLVSHSSYFVPEPSDLQSAVSIILATPQS 184  
Db 141 NEGSSVTLMCLAFGRPE-----PTVTR-----HLSGKGQGFVSEDEYLE-----ITGITREQ 188  
QY 185 NGTITCVATWKSLSKARKSATVNLTV-----IRCQDTGGGINIPGVL-----SSLPSLGFS 235  
Db 189 SGEYCSAV-NDVAVPDVRKVTVNPYPPTISNAKNTGASVGQKGLQCEASAVPVAEFQ 247  
QY 236 LPTWKGVLGLA 247  
Db 248 ---WFKEDTRLA 256  
RESULT 3  
NTRI RAT  
ID NTRI RAT STANDARD; PRT; 344 AA.  
AC Q62718;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Neurotrophin precursor (GP65).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 217-229.  
RC STRAIN=Sprague-Dawley;  
RA MEDLINE=95198094; PubMed=7891157;  
RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,  
RA Salzer J.L.;  
RT "Cloning of neurotrophin defines a new subfamily of differentially  
RT expressed neural cell adhesion molecules.";  
RL J. Neurosci. 15:2141-2156(1995).  
CC -!- FUNCTION: Neural cell adhesion molecule.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN SEVERAL  
CC SUBPLATE, AND LOWER CORTICAL LAMINAE IN THE FOREBRAIN AND IN THE  
CC PONTINE NUCLEUS, CEREBELLAR GRANULE CELLS, AND PURKINJE CELLS IN  
CC THE HINDBRAIN.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U16845; AAA67445.1; -.  
CC PIR; I56551; I56551.  
CC InterPro; IPR007110; IG-like.  
CC InterPro; IPR003598; IG\_C2.  
CC InterPro; IPR003006; IG\_MHC.  
CC Pfam; PF00047; Ig; 3.  
CC SMART; SM00408; IGC2; 2.  
CC PROSITE; PS00835; IG LIKE; 3.  
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
CC Repeat; Signal.  
CC POTENTIAL.  
CC CHAIN 31 321  
CC PROPEP 322 344  
CC DOMAIN 39 126  
CC DOMAIN 136 218  
CC DOMAIN 222 309  
CC DISULFID 57 115  
FT SIGNAL 1 31  
FT CHAIN 32 321  
FT PROPEP 322 344  
FT DOMAIN 39 126  
FT DOMAIN 136 218  
FT DOMAIN 222 309  
FT DISULFID 57 115

```

FT DISULFID 157 201 POTENTIAL.
FT DISULFID 243 295 POTENTIAL.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 321 321 GPI-ANCHOR (POTENTIAL).
SQ SEQUENCE 344 AA; 37998 MW; CBB39BE5B3B324 CRC64;

Query Match 6.3%; Score 129.5; DB 1; Length 344;
Best Local Similarity 24.2%; Pred. No. 0.0039;
Matches 64; Conservative 43; Mismatches 104; Indels 53; Gaps 14;

QY 4 VIFLHSGS-----SGNEVI-EGPQNAVTLKGSQARFNCTVSQGWKLIMWALSDMVVLS---55
Db 21 LLELVPTGVPRSGDATFPKAMDVNTVRQGESATLRCITDNRVTRVAVLNRSTILYAGND 80
QY 56 ---VRPMEPIITNDRTSQDYDQGNFTSEMIIHNVEPDSGNIRCSLQ-----NSRLH 106
Db 81 KWCLDPRVLLSN---TQTOY-----SIBQNVVDYDEGPTCSVDNDHPKTSRVH 129
QY 107 GSAYLTVQMGELFIPSVNLVVAENEPCEVTCLPSSHWRTPDISWELGLLVSHSSYFFVP 166
Db 130 ----LIVQSPKIVEISSDISINEGNNISLTCIATGRPE-PTVTWR---HISPKAVGFVS 181
QY 167 EPDLQSAVILALTPQSGNTLFCVATKSLKARKSATVNLTVIRCP-----QDTGGGIN 221
Db 182 EDEYLE-----IQGITRQSGEYECSSA-NDVAAPVVRVNTVNTVPYIYSEAKGTGPVG 236
QY 222 IPGVL-----SSLPSLGSFLTGWK 241
Db 237 QKGLTQCEASAVPSAEFQ---WFK 257

RESULT 4
CEPU_CHICK STANDARD; PRT; 353 AA.
ID_CEPU_CHICK
AC Q90773;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CEPU-1 protein precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RA MEDLINE=96370549; PubMed=8774445;
RX Spaltmann F., Bruemmendorf T.;
RT developing cerebellar Purkinje cells.;
RL J. Neurosci. 16:1770-1779(1996).
CC -!- FUNCTION: IT MAY BE A CELLULAR ADDRESS MOLECULE SPECIFIC TO
CC PURKINJE CELLS. IT MAY REPRESENT A RECEPTOR OR A SUBUNIT OF A
CC RECEPTOR COMPLEX.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Minor;
CC IsoId=Q90773-1; Sequence=Displayed;
CC Name=2; Synonyms=Major;
CC IsoId=Q90773-2; Sequence=VSP_002607;
CC -!- TISSUE SPECIFICITY: FOUND ON THE DENDRITES, SOMATA AND AXONS OF
CC DEVELOPING PURKINJE CELLS. UNDETECTABLE ON OTHER NEURONS LIKE
CC GOLGI OR GRANULE CELLS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED BY DEVELOPING CEREBELLAR PURKINJE

```

```

CC CELLS. EXPRESSION COINCIDES WITH THE GROWTH OF THE DENDRITIC TREE,
CC AFTER PURKINJE CELLS HAVE FINISHED THEIR MIGRATION FROM THE
CC VENTRICULAR ZONE (FROM E15 UNTIL E21). EXPRESSED IN THE ADULT.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z72497; CAA96578.1; -.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00408; Igc2; 2.
CC PROSITE; PS0835; IGLIKE; 3.
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
CC Repeat; Signal; Alternative splicing
CC SIGNAL 1 28
CC CHAIN 29 330 CEPU-1 PROTEIN.
CC PROPEP 331 353 REMOVED IN NATURE FORM (POTENTIAL).
CC DOMAIN 37 124 IG-LIKE C2-TYPE 1.
CC DOMAIN 134 216 IG-LIKE C2-TYPE 2.
CC DOMAIN 220 314 IG-LIKE C2-TYPE 3.
CC DISULFID 55 113 POTENTIAL.
CC DISULFID 155 199 POTENTIAL.
CC DISULFID 241 293 POTENTIAL.
CC CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC LIPID 330 330 GPI-ANCHOR (POTENTIAL).
CC VARSPLIC 310 320 Missing (in isoform 2).
CC /FTId=VSP_002607.
CC SQ SEQUENCE 353 AA; 38736 MW; 2550C48591EBBA6 CRC64;

Query Match 6.2%; Score 128; DB 1; Length 353;
Best Local Similarity 25.7%; Pred. No. 0.0054;
Matches 63; Conservative 36; Mismatches 90; Indels 56; Gaps 13;

QY 22 NATVLKGSQARFNCTYSQ-----GW---KLIMWALSDMVVLSVRPMEPIITNDRTSQRY 73
Db 42 NVTVRQGESATLRCSDVNDNRVTRVAVLNRSSILYAGNDKWCLDPRVLLANTKTOYSIQ-- 99
QY 74 DQGNFTSEMIHNVEPDSGNIRCSLQ-----NSRLHGSAYLTVQMGELFIPSVNLV 127
Db 100 -----IHDVDYDEGPTCSVDNDHPKTSRVH---LIVQSPKITETSSDIS 144
QY 128 VAENPECVTCLPSHWRTPD--ISWELGLLVSHSSYFFVPSPDLSOASVIALTPQSN 185
Db 145 INEGGNVSLTCIA---TGRPDPTITWR---HISPKAVGFISEDEYLE---ITGITREQS 194
QY 186 GTTTCVATWKSLEKARKSATVNLTV-----IRCPQDTGGGINIPQVL-----SSLPSLGSFL 236
Db 195 GEVECSAS-NDVAAPVVQRVKVTNVPYIYSDAKSTGTVPGVQXGILMCEASAVPSADFQ- 252
QY 237 PTWKGK 241
Db 253 --WYK 255

RESULT 5
DCC_HUMAN
ID_DCC_HUMAN STANDARD; PRT; 1447 AA.
AC P43146;

```

DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Tumor suppressor protein DCC precursor (Colorectal cancer suppressor).  
 GN DCC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95011532; PubMed=7926722;  
 RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,  
 RA Vogelstein B.;  
 RT "The DCC gene product in cellular differentiation and colorectal  
 RT tumorigenesis.";  
 RL Genes Dev. 8:1174-1183 (1994).  
 RN [2]  
 RN SEQUENCE OF 1-750 FROM N.A.  
 RP MEDLINE=90100559; PubMed=2294591;  
 RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,  
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,  
 RA Vogelstein B.;  
 RT "Identification of a chromosome 18q gene that is altered in  
 RT colorectal cancers.";  
 RL Science 247:49-56 (1990).  
 RN [3]  
 RN SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).  
 RP MEDLINE=91121517; PubMed=1991322;  
 RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,  
 RA Oliner J.D., Kinzler K.W., Vogelstein B.;  
 RT "Scrambled exons.";  
 RL Cell 64:607-613 (1991).  
 RN [4]  
 RN GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.  
 RP MEDLINE=94245241; PubMed=8188295;  
 RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,  
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;  
 RT "The DCC gene: structural analysis and mutations in colorectal  
 RT carcinomas.";  
 RL Genomics 19:525-531 (1994).  
 RN [5]  
 RN VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.  
 RP MEDLINE=94243823; PubMed=8187090;  
 RA Miyake S., Nagai K., Yoshino K., Oto M., Endo M., Yuasa Y.;  
 RA "Point mutations and allelic deletion of tumor suppressor gene DCC in  
 RA human esophageal squamous cell carcinomas and their relation to  
 RA metastasis.";  
 RL Cancer Res. 54:3007-3010 (1994).  
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL  
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.  
 CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO  
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC  
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR  
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS  
 CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 6 fibronectin type III domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; X76132; CAA53735.1; -;  
 DR EMBL; M32292; AAA35751.1; -;

DR EMBL; M32286; AAA52174.1; -;  
 DR EMBL; M32288; AAA52175.1; ALT\_SEQ.  
 DR EMBL; M32290; AAA52176.1; -;  
 DR EMBL; M63696; AAA52177.1; -;  
 DR EMBL; M63700; AAA52178.1; -;  
 DR EMBL; M63702; AAA52179.1; -;  
 DR EMBL; M63718; AAA52180.1; -;  
 DR EMBL; M63698; AAA52181.1; -;  
 DR PIR; A54100; A54100.  
 DR HSP; P56276; ITLK.  
 DR Genew; HGNC:2701; DCC.  
 DR MIM; 120470; -;  
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO; GO:0007409; P:axonogenesis; TAS.  
 DR GO; GO:0006917; P:induction of apoptosis; TAS.  
 DR GO; GO:0007048; P:oncogenesis; TAS.  
 DR InterPro; IPR003961; FN.III.  
 DR InterPro; IPR003962; FN.III.subd.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig C2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00041; fn3; 6.  
 DR Pfam; PF00047; ig; 4.  
 DR PRINTS; PR00014; FN1PEIII.  
 DR SMART; SM00060; FN3; 6.  
 DR SMART; SM00408; IGC2; 3.  
 DR PROSITE; PSS0835; IG LIKE; 4.  
 DR Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;  
 KW Repeat; Anti-oncogene; Disease mutation; Polymorphism.  
 FT SIGNAL 1 25  
 FT CHAIN 26 1447  
 FT DOMAIN 26 1097  
 FT TRANSMEM 1098 1122  
 FT DOMAIN 1123 1447  
 FT DOMAIN 26 135  
 FT DOMAIN 139 229  
 FT DOMAIN 234 326  
 FT DOMAIN 331 416  
 FT DOMAIN 426 522  
 FT DOMAIN 525 618  
 FT DOMAIN 619 716  
 FT DOMAIN 722 816  
 FT DOMAIN 840 940  
 FT DOMAIN 941 1042  
 FT DISULFID 61 117  
 FT DISULFID 161 212  
 FT DISULFID 261 310  
 FT DISULFID 352 400  
 FT CARBOHYD 94 94  
 FT CARBOHYD 299 299  
 FT CARBOHYD 318 318  
 FT CARBOHYD 478 478  
 FT CARBOHYD 628 628  
 FT CARBOHYD 702 702  
 FT VARIANT 168 168  
 FT VARIANT 201 201  
 FT VARIANT 1375 1375  
 FT CONFLICT 138 138  
 FT CONFLICT 233 329  
 FT CONFLICT 421 421  
 SQ SEQUENCE 1447 AA; 158456 MW; 4A8612766ED0471F CRC64;  
 Query Match 6.28; Score 128; DB 1; Length 1447;  
 Best Local Similarity 24.8%; Pred. No. 0.031;  
 Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;  
 QY 10 SGSGNEV-----IEGPQATVLKSGQARFNCTVSQGW--KLIMWALSDM 51  
 Db 220 SRIGNEAEVILSDPGLHRLQYFLQPSNVVAIEGDAVLECCVS-GYPPSPFWLGRGE 278

QY 52 VVLSVRPMEPIITNDRTSQRDQGNFTSEMIHNVPSDSNIRC--SLQNSRLHGSA 109  
 Db 279 VI-----QLRSKSLG--SNLISNVTDGSMVTCVVTYKKNISASA 323  
 QY 110 YLTQVQMGELFIPSVNLVVAENPCVETCLPSPHWTLPDISM-ELGLLYSHSYFVPEP 168  
 Db 324 ELTVLPVPFLNPSNLAYESMDIEFCTVS-GKVPETVNMKNGDVIPSDYFQIVGG 382  
 QY 169 SDLOSASVIALTPQNGHILTCVATWKSLSKARSATVNLTVIRCPDGTGGINIPGVLS 228  
 Db 383 SNLR-----ILGVVKSDEGFYQCAENAGNAQTSQALIVKPKAIPSS-----VLPS 430  
 QY 229 LP 230  
 Db 431 AP 432  
 RESULT 6  
 KMLS HUMAN STANDARD; PRT; 1914 AA.  
 AC Q15746; O95796; O95797; O95798; O95799; Q14844; Q16794; Q9UBG5;  
 AC Q9UIT9;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Myosin light chain kinase, smooth muscle and non-muscle isoforms  
 DE (EC 2.7.1.117) (MLCK) [Contains: Telokin (Kinase related protein)  
 DE (KRP)].  
 GN MYLK OR MLCK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Umbilical vein endothelial cells;  
 RX MEDLINE=97304466; PubMed=9160829;  
 RA Garcia J.G.N., Lazar V.L., Gilbert-McClain L.I., Gallagher P.J.,  
 RA Varin A.D.;  
 RT "Myosin light chain kinase in endothelium: molecular cloning and  
 RT regulation.";  
 RL Am. J. Respir. Cell Mol. Biol. 16:489-494 (1997).  
 RN [2]  
 RP REVISIONS.  
 RA Burkov K.G., Garcia J.G.N.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2; 3A; 3B AND 4).  
 RC TISSUE=Umbilical vein;  
 RX MEDLINE=99216419; PubMed=10198165;  
 RA Lazar V.L., Garcia J.G.N.;  
 RT "A single human myosin light chain kinase gene (MLCK; MYLK).";  
 RL Genomics 57:256-267 (1999).  
 RN [4]  
 RP REVISIONS (ISOFORM 2).  
 RA Burkov K.G., Garcia J.G.N.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 923-1914 FROM N.A.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=96121365; PubMed=8575746;  
 RA Potier M.-C., Chelot E., Fekarsky Y., Gardiner K., Rossier J.,  
 RA Turnell W.G.;  
 RT "The human myosin light chain kinase (MLCK) from hippocampus:  
 RT cloning, sequencing, expression, and localization to 3qcen-q21.";  
 RL Genomics 29:562-570 (1995).  
 RN [6]  
 RP SEQUENCE OF 1614-1914 FROM N.A.  
 RC TISSUE=Lung, and Placenta;  
 RX MEDLINE=20007838; PubMed=10536370;  
 RA Watterson D.M., Schavocky J.P., Guo L., Weiss C., Chlenski A.,  
 RA Shinsky V.P., Van Eldik L.J., Haiech J.;  
 RT "Analysis of the kinase-related protein gene found at human chromosome

3q21 in a multi-gene cluster: organization, expression, alternative  
 splicing and polymorphic marker.";  
 J. Cell. Biochem. 75:481-491 (1999).  
 [7]  
 RP SEQUENCE OF 1456-1914 FROM N.A.  
 RC TISSUE=Placenta;  
 RA Watterson M.D.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR  
 CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE  
 CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT  
 CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS. CENTRAL  
 CC DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE  
 CC EDEMA FORMATION. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO  
 CC CONTROL THE GROWTH INITIATION OF ASTROCYTIC PROCESSES IN CULTURE  
 CC AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED  
 CC BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT  
 CC IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.  
 CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin  
 CC light-chain] phosphate.  
 CC -!- SUBUNIT: ISOFORM TELOKIN BINDS CALMODULIN.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=6;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1; Synonyms=Non-muscle isoform;  
 CC IsoId=Q15746-1; Sequence=Displayed;  
 CC Note=The smooth muscle isoform and telokin are produced by  
 CC alternative initiation at Met-923 and Met-1761 of isoform 1;  
 CC Name=2;  
 CC IsoId=Q15746-2; Sequence=VSP\_004791;  
 CC Name=3A;  
 CC IsoId=Q15746-3; Sequence=VSP\_004792, VSP\_004794;  
 CC Name=3B;  
 CC IsoId=Q15746-4; Sequence=VSP\_004791, VSP\_004792, VSP\_004794;  
 CC Name=4;  
 CC IsoId=Q15746-5; Sequence=VSP\_004792, VSP\_004793;  
 CC Name=Del-1790;  
 CC IsoId=Q15746-6; Sequence=VSP\_004795;  
 CC Event=Alternative initiation;  
 CC Comment=3 isoforms, 1/Non-muscle isoform (shown here);  
 CC smooth-muscle isoform and telokin, are produced by alternative  
 CC initiation at Met-1, Met-923 and Met-1761. Telokin has no  
 CC catalytic activity;  
 CC -!- TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOFORMS ARE  
 CC EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN  
 CC CULTURED ENDOTHELIUM WITH QUALITATIVE EXPRESSION APPEARING TO BE  
 CC NEITHER TISSUE- NOR DEVELOPMENT-SPECIFIC. NON-MUSCLE ISOFORM 2 IS  
 CC THE DOMINANT SPLICE VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN  
 CC HAS BEEN FOUND IN A WIDE VARIETY OF ADULT AND FETAL TISSUES.  
 CC -!- PTM: MLCK IS PROBABLY DOWN-REGULATED BY PHOSPHORYLATION.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -!- SIMILARITY: Contains 9 immunoglobulin-like C2-type domains.  
 CC -!- SIMILARITY: Contains 1 fibronectin type III domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U48959; AAC18423.2; -;  
 CC EMBL; AF069601; AAD15921.2; -;  
 CC EMBL; AF069602; AAD15922.1; -;  
 CC EMBL; AF069603; AAD15923.1; -;  
 CC EMBL; AF069604; AAD15924.1; -;  
 CC EMBL; X85337; CAA59685.1; -;  
 CC EMBL; AF096771; AAD51380.1; -;  
 CC EMBL; AF096766; AAD51380.1; JOINED.  
 CC EMBL; AF096767; AAD51380.1; JOINED.  
 CC EMBL; AF096768; AAD51380.1; JOINED.  
 CC EMBL; AF096769; AAD51380.1; JOINED.

DR EMBL; AF096770; AAD51380.1; JOINED.  
DR EMBL; AF096774; AAD54018.1; -.  
DR EMBL; AF096771; AAD51381.1; -.  
DR EMBL; AF096769; AAD51381.1; JOINED.  
DR EMBL; AF096770; AAD51381.1; JOINED.  
DR EMBL; AF096770; AAD51381.1; JOINED.  
DR HSP; Q63450; 1A06.  
DR Genew; HGNC:7590; MYLK.  
DR MIM; 600922; -.  
DR GO; GO:0004687; F-Myosin light chain kinase activity; TAS.  
DR GO; GO:0004688; P-protein amino acid phosphorylation; TAS.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR Pfam; PF00041; fn3; 1.  
DR Pfam; PF00047; Ig; 8.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Prot kinase; 1.  
DR PROSITE; PS00835; IG LIKE; 9.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;  
KW ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;  
KW Alternative initiation; Alternative splicing.  
CHAIN 1 1914  
FT CHAIN 923 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM NON-  
FT CHAIN 1761 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM  
SMOOTH-MUSCLE ISOZYME.  
FT CHAIN 923 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM  
SMOOTH-MUSCLE ISOZYME.  
FT CHAIN 1761 1914 MYOSIN LIGHT CHAIN KINASE, ISOFORM  
TELOKIN.  
FT INIT\_MET 923 923 FOR ISOFORM SMOOTH-MUSCLE ISOZYME.  
FT INIT\_MET 1761 1761 FOR ISOFORM TELOKIN.  
FT DOMAIN 33 122 IG-LIKE C2-TYPE 1.  
FT DOMAIN 161 249 IG-LIKE C2-TYPE 2.  
FT DOMAIN 414 503 IG-LIKE C2-TYPE 3.  
FT DOMAIN 514 599 IG-LIKE C2-TYPE 4.  
FT DOMAIN 620 711 IG-LIKE C2-TYPE 5.  
FT DOMAIN 721 821 IG-LIKE C2-TYPE 6.  
FT DOMAIN 1098 1186 IG-LIKE C2-TYPE 7.  
FT DOMAIN 1238 1326 IG-LIKE C2-TYPE 8.  
FT DOMAIN 1343 1413 FIBRONECTIN TYPE-III.  
FT DOMAIN 1464 1719 PROTEIN KINASE.  
FT DOMAIN 1711 1774 CALMODULIN-BINDING.  
FT DOMAIN 1809 1898 IG-LIKE C2-TYPE 9.  
FT NP\_BIND 1470 1478 ATP (BY SIMILARITY).  
FT BINDING 1493 1493 ATP (BY SIMILARITY).  
FT ACT\_SITE 1585 1585 BY SIMILARITY.  
FT DOMAIN 1906 1914 POLY-GLU.  
FT DOMAIN 868 998 5 X 28 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 868 895 I-1.  
FT REPEAT 895 923 I-2.  
FT REPEAT 924 951 I-3.  
FT REPEAT 952 979 I-4.  
FT REPEAT 980 998 I-5 (INCOMPLETE).  
FT DOMAIN 999 1063 6 X 12 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 999 1003 II-1 (INCOMPLETE).  
FT REPEAT 1004 1015 II-2.  
FT REPEAT 1016 1027 II-3.  
FT REPEAT 1028 1039 II-4.  
FT REPEAT 1040 1051 II-5.  
FT REPEAT 1052 1063 II-6.  
FT VARSPLIC 437 506 VSGIPKPEVAMFLEGTVPVROGSEIVVEDAGSHYLLKA  
FT RTRDSGTVCSTASNAQGVCSWTQVVER -> G (in  
FT isoform 2 and isoform 3B).  
FT /FTid=VSP\_004791.  
FT DEVEVSD -> MKWRCCQT (in isoform 3A,  
FT isoform 3B and isoform 4).  
FT /FTid=VSP\_004792.  
FT Missing (in isoform 4).  
FT /FTid=VSP\_004793.

Query Match 6.2%; Score 128; DB 1; Length 1914;  
Best Local Similarity 20.3%; Pred. No. 0.044;  
Matches 86; Conservative 58; Mismatches 149; Indels 130; Gaps 19;  
QY 20 PQNATVLKGSQARFNCTVSGW--KLIMWALSVMVLVSVPMBPIIINDRFTSRYDQGG 77  
DB 39 PRNLCKEGATAKFRGRV-RGYPEQVTH-----RNGQPITSGGRFL---LDCGI 85  
QY 78 NFTSEMIHNVEPSDSGNIRCSLQNSRLHGSAYLTVOV-----MGEI 119  
DB 86 RGTFSLVIHVAHEEDRGKYTCBATNG--SGARQVVELTVGSGFAKQLGQPVVSKTLGDR 143  
QY 120 FI-----PSV-----NLVVAENE-----PCEVTCCLPSHWTLPLDISWELG- 154  
DB 144 FSASAVETREPSIWGECPPKFPATKLGVRVVKVGQMGREFSKITGRPQ-----PQVTLKGN 198  
QY 155 LLYSHSSYFVPEPSDLOSASVILALTPQSNGLTLCVATWKSILKAKSA-----TVN 206  
DB 199 VPLQPSARVSVSEKNGMQ-VLEIHGVNDVDGVYVTVCLVNVGSGKASMSBELSIQGLDSAN 257  
QY 207 LTWIRCPQDTGGGI-----NIPGVLSLPSLGFSLPTWGVGLAGTLLTPTCTLTIR 261  
DB 258 RSFVRETKATNSDVREKVTNIVSKESKLSL----- 288  
QY 262 CCCRRRCGCCNCCRCFCRRKRGFRIOFKKSEKKTNETETESGNEN---SGVNS 318  
DB 289 -----EAAAKSKNCSSPQGGSPPAWANSQPQP-RESKLESCKDSPTAPQTP 336  
QY 319 DEQKTTDTASLPPKSCSSDPQBNSSCGPHQADOR-----PPRASHQOASNLASPE 374  
DB 337 VLQKTSSTITQAARVQ---PEPRAPGLVLSPEGEERKPPRPPATFTTROPGLSQSD 393  
QY 375 KVS 377  
DB 394 VVS 396  
RESULT 7  
ID\_DCC\_MOUSE STANDARD; PRT; 1447 AA.  
AC P70211;1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tumor suppressor protein DCC precursor.  
GN DCC  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).  
RC STRAIN=BALB/c; TISSUE=Brain;  
RX MEDLINE=96112625; Pubmed=8570174;  
RA Cooper H.M., Armes P., Britto J., Gad J., Wilks A.F.;  
RT \*Cloning of the mouse homologue of the deleted in colorectal cancer  
RL gene (mDCC) and its expression in the developing mouse embryo.;  
RN Oncogene 11:2243-2254(1995).  
RP REVISIONS.  
RC STRAIN=BALB/c; TISSUE=Brain;  
RA Cooper H.M.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=A;  
CC IsoId=P70211-1; Sequence=Displayed;  
CC Note=Isoform B is produced by alternative initiation at Met-85  
CC of isoform A;  
CC Name=C;

CC		Isoid=P70211-2; Sequence=VSP_002501;
CC	Event=Alternative initiation;	
CC	Comment=2 isoforms, A (shown here) and B, are produced by	
CC	alternative initiation at Met-1 and Met-85;	
CC	-!- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE	
CC	DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN	
CC	BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.	
CC	ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.	
CC	-!- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS	
CC	EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION	
CC	AND REMAIN AT THIS LEVEL IN THE ADULT.	
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC	
CC	SUPERFAMILY.	
CC	-!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.	
CC	-!- SIMILARITY: Contains 6 fibronectin type III domains.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to licens@isb-sib.ch).	
CC	-----	
DR	EMBL; X85788; CAA59786.1; --	
DR	HSSP; P56276; ITLX	
DR	MGD; MGJ:94869; DCC	
DR	InterPro; IPR003961; FN.III.	
DR	InterPro; IPR003962; FN.III subd.	
DR	InterPro; IPR007110; Ig-Like	
DR	InterPro; IPR003598; Ig_C2	
DR	InterPro; IPR003006; Ig_MHC	
DR	Pfam; PF00041; fn3; 6	
DR	Pfam; PF00047; ig; 4	
DR	PRINTS; PR00014; FNTYPEIII	
DR	SMART; SM00060; FN3; 6	
DR	SMART; SM00408; IGC2; 3	
DR	PROSITE; PS50835; IG_LIKE; 4	
DR	Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;	
KW	Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.	
KW	GLYCOPROTEIN; IMMUNOGLOBULIN DOMAIN; TRANSMEMBRANE; SIGNAL;	
FT	SIGNAL 1 25 POTENTIAL	
FT	CHAIN 26 1447 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM A.	
FT	CHAIN 85 1447 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM B.	
FT	INIT MET 85 85 FOR ISOFORM B.	
FT	DOMAIN 26 1097 EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 1098 1122 CYTOPLASMIC (POTENTIAL).	
FT	DOMAIN 1123 1447	
FT	DOMAIN 36 135 IG-LIKE C2-TYPE 1.	
FT	DOMAIN 139 229 IG-LIKE C2-TYPE 2.	
FT	DOMAIN 234 326 IG-LIKE C2-TYPE 3.	
FT	DOMAIN 331 416 IG-LIKE C2-TYPE 4.	
FT	DOMAIN 426 522 FIBRONECTIN TYPE-III 1.	
FT	DOMAIN 525 618 FIBRONECTIN TYPE-III 2.	
FT	DOMAIN 619 716 FIBRONECTIN TYPE-III 3.	
FT	DOMAIN 722 816 FIBRONECTIN TYPE-III 4.	
FT	DOMAIN 840 940 FIBRONECTIN TYPE-III 5.	
FT	DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.	
FT	DISULFID 61 117 BY SIMILARITY.	
FT	DISULFID 161 212 BY SIMILARITY.	
FT	DISULFID 261 310 BY SIMILARITY.	
FT	DISULFID 352 400 BY SIMILARITY.	
FT	CARBOHYD 60 60 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	CARBOHYD 94 94 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	CARBOHYD 299 299 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	CARBOHYD 318 318 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	CARBOHYD 478 478 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	CARBOHYD 628 628 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	CARBOHYD 702 702 N-LINKED (GLCNAC. . ) (POTENTIAL).	
FT	VARSPPLIC 819 838 Missing (in isoform C).	
FT	/FTID=VSP_002501.	
SQL	SEQUENCE 1447 AA; 158298 MW; 0D1F1097C2D5B9F CRC64;	

### Query Match

```

Best Local Similarity 26.6%; Pred. No. 0.037;
Matches 51; Conservative 29; Mismatches 86; Indels 26; Gaps 8;

QY 17 IEQPQATVLKGSQARFNCTVSQW--KLIMWALSDMVVLVSRPMEPIITNDRFTSQRYD 74
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
DB 243 LQRPNSVIALEGKDAVLECCVS-GYPFPSTWLRGEEVI-----QLRSKKYS 288
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

QY 75 QQGNFTSEMIIHVSPDSGNTRC--SLQNSRLHGSAYLTQVMGELFIPSVNVLVAENE 132
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
DB 289 LLGG--SNLLISNVTDDSGTYTCVVTYKNETISASAEULTVLPFPWFLNHPNSLYAYESM 346
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

QY 133 PCGVTCILPSHWTRLPDISW-ELGLLVSHSYYVPVPPSDLOSASVLSIALTPQSGNGLTCV 191
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
DB 347 DIFECQVS-GKVPPTVNMKNGDVIPSDYFQIVGGSNLR-----ILGVVKSDEGFQCV 401
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

QY 192 ATWKSLLKARKSA 203
   | :: :: :: ::
DB 402 AENEAGNAQSSA 413

RESULT 8
UN89 CAEEL
ID UN89 CAEEL STANDARD; PRT; 6632 AA.
AC O01761; Q17362;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
DB NCBI_TaxID=6239;
GN UNC-89 OR C09D1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Bristol N2;
RX MEDLINE=96130278; PubMed=8603916;
RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;
RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
RT assembly, encodes a giant modular protein composed of Ig and signal
RT transduction domains".
RT J. Cell Biol. 132:835-848 (1996).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Du Z., Le T.T., Wilson R.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Structural component of the muscle M-line. Myofilament
CC lattice assembly begins with positional cues laid down in the
CC basement membrane and muscle cell membrane. UNC-89 responds to
CC these signals, localizes, and then participates in assembling an
CC M-line.
CC -!- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 5 RCSD domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U33058; AAB00542.1; --

```

```

DR EMBL; AF003131; AAB54132.2; -.
DR PDB; 1FHO; 20-DEC-00.
DR WormPep; C09D1.1; CE30426.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001849; PH.
DR InterPro; IPR007850; RCHD.
DR InterPro; IPR000219; RhoGEF.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; Ig; 47.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF05177; RCHD; 5.
DR Pfam; PF00621; RhoGEF; 1.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00408; IGC2; 23.
DR SMART; SM00325; RhoGEF; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS0010; DH 2; 1.
DR PROSITE; PS50835; IG LIKE; 49.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50002; SH3; 1.
KW Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
KW 3D-structure.
FT DOMAIN 63 127 SH3.
FT DOMAIN 152 330 DH.
FT DOMAIN 342 498 PH.
FT DOMAIN 547 633 IG-LIKE C2-TYPE 1.
FT DOMAIN 548 736 IG-LIKE C2-TYPE 2.
FT DOMAIN 748 838 IG-LIKE C2-TYPE 3.
FT DOMAIN 946 1033 IG-LIKE C2-TYPE 4.
FT DOMAIN 1044 1132 IG-LIKE C2-TYPE 5.
FT DOMAIN 1140 1227 IG-LIKE C2-TYPE 6.
FT DOMAIN 1272 1315 THR-RICH.
FT DOMAIN 1375 1475 RCHD 1.
FT DOMAIN 1479 1585 RCHD 2.
FT DOMAIN 1597 1695 RCHD 3.
FT DOMAIN 1700 1799 RCHD 4.
FT DOMAIN 1800 1860 RCHD 5.
FT DOMAIN 1982 2067 IG-LIKE C2-TYPE 7.
FT DOMAIN 2071 2163 IG-LIKE C2-TYPE 8.
FT DOMAIN 2171 2261 IG-LIKE C2-TYPE 9.
FT DOMAIN 2269 2359 IG-LIKE C2-TYPE 10.
FT DOMAIN 2367 2455 IG-LIKE C2-TYPE 11.
FT DOMAIN 2463 2564 IG-LIKE C2-TYPE 12.
FT DOMAIN 2563 2651 IG-LIKE C2-TYPE 13.
FT DOMAIN 2657 2746 IG-LIKE C2-TYPE 14.
FT DOMAIN 2754 2858 IG-LIKE C2-TYPE 15.
FT DOMAIN 2887 2980 IG-LIKE C2-TYPE 16.
FT DOMAIN 2994 3081 IG-LIKE C2-TYPE 17.
FT DOMAIN 3087 3183 IG-LIKE C2-TYPE 18.
FT DOMAIN 3189 3280 IG-LIKE C2-TYPE 19.
FT DOMAIN 3286 3376 IG-LIKE C2-TYPE 20.
FT DOMAIN 3384 3469 IG-LIKE C2-TYPE 21.
FT DOMAIN 3482 3572 IG-LIKE C2-TYPE 22.
FT DOMAIN 3580 3667 IG-LIKE C2-TYPE 23.
FT DOMAIN 3677 3777 IG-LIKE C2-TYPE 24.
FT DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.
FT DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.
FT DOMAIN 4018 4106 IG-LIKE C2-TYPE 27.
FT DOMAIN 4201 4297 IG-LIKE C2-TYPE 28.
FT DOMAIN 4312 4397 IG-LIKE C2-TYPE 29.
FT DOMAIN 4400 4485 IG-LIKE C2-TYPE 30.
FT DOMAIN 4489 4580 IG-LIKE C2-TYPE 31.
FT DOMAIN 4588 4678 IG-LIKE C2-TYPE 32.
FT DOMAIN 4681 4771 IG-LIKE C2-TYPE 33.
FT DOMAIN 4873 4961 IG-LIKE C2-TYPE 34.
FT DOMAIN 4965 5057 IG-LIKE C2-TYPE 35.
FT DOMAIN 5067 5160 IG-LIKE C2-TYPE 36.
FT DOMAIN 5171 5260 IG-LIKE C2-TYPE 37.
FT DOMAIN 5260 5360 IG-LIKE C2-TYPE 38.
FT DOMAIN 5366 5472 IG-LIKE C2-TYPE 39.
FT DOMAIN 5472 5578 IG-LIKE C2-TYPE 40.
FT DOMAIN 5578 5685 IG-LIKE C2-TYPE 41.
FT DOMAIN 5685 5790 IG-LIKE C2-TYPE 42.
FT DOMAIN 5790 5904 IG-LIKE C2-TYPE 43.
FT DOMAIN 5904 6014 IG-LIKE C2-TYPE 44.
FT DOMAIN 6014 6130 IG-LIKE C2-TYPE 45.
FT DOMAIN 6130 6239 IG-LIKE C2-TYPE 46.
FT DOMAIN 6239 6368 IG-LIKE C2-TYPE 47.
FT DOMAIN 6368 6502 FIBROECTIN TYPE-III.
FT DOMAIN 6502 6596 IG-LIKE C2-TYPE 48.
FT DOMAIN 6596 6721 IG-LIKE C2-TYPE 49.
FT DISULFID 568 621 POTENTIAL.
FT DISULFID 2908 2975 POTENTIAL.
FT DISULFID 3015 3065 POTENTIAL.
FT DISULFID 3707 3759 POTENTIAL.
FT DISULFID 3826 3890 POTENTIAL.
FT DISULFID 5092 5157 POTENTIAL.
FT DISULFID 5298 5350 POTENTIAL.
FT DISULFID 5508 5560 POTENTIAL.
FT DISULFID 5616 5669 POTENTIAL.
FT DISULFID 5722 5764 POTENTIAL.
FT DISULFID 5836 5901 POTENTIAL.
FT DISULFID 5946 5998 POTENTIAL.
FT DISULFID 6036 6171 POTENTIAL.
FT DISULFID 6421 6486 POTENTIAL.
FT CONFLICT 2137 2137 A -> P (IN REF. 1).
FT CONFLICT 2245 2247 AKA -> PKP (IN REF. 1).
FT CONFLICT 2258 2258 A -> P (IN REF. 1).
FT CONFLICT 2284 2284 E -> G (IN REF. 1).
FT CONFLICT 2297 2297 M -> I (IN REF. 1).
FT CONFLICT 3531 3531 A -> G (IN REF. 1).
FT CONFLICT 3884 3888 DAGEY -> RRRRI (IN REF. 1).
FT CONFLICT 3929 3929 A -> V (IN REF. 1).
FT CONFLICT 5134 5134 A -> P (IN REF. 1).
FT CONFLICT 5145 5145 T -> S (IN REF. 1).
FT CONFLICT 5185 5185 G -> A (IN REF. 1).
FT CONFLICT 5199 5199 K -> N (IN REF. 1).
FT CONFLICT 5202 5202 L -> F (IN REF. 1).
FT CONFLICT 5213 5213 F -> L (IN REF. 1).
FT CONFLICT 6178 6178 A -> G (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SQ SEQUENCE 6632 AA; 731665 MW; 262D3EDD62960E89 CRC64;
Query Match 6.1%; Score 125.5; DB 1; Length 6632;
Best Local Similarity 23.9%; Pred. No. 0.33;
Matches 52; Conservative 39; Mismatches 92; Indels 35; Gaps 8;
Qy 16 VIEGPQNAVTLKGSQARFNCTVSQGW--KLIMWALS DMVLSVRPMEPIIINDRETSQRY 73
Db 2077 VVDGPKSVTIKETETAEFKATIS-GFPAPT VKWTINEKIVEESRTITTIKTEDVYI---- 2131
Qy 74 DQGNFTSEMIHNVEPSDSGNIRCSLQNSRLHGSAYLTQVMGELFIPS-----VNLVV 128
Db 2132 -----LKISNAKIEQTGTVKVTAQNSAGQSKQADLKVENVKAPKESQITDKVA 2182
Qy 129 AENEPCEVTCL---PSHWRLPDISWEL-GLLVSHSYFYFVPEPSDLQSAVSIILALTPQS 184
Db 2183 DEGEPLRWNLDELDPSPGT---EVSWLINGQPLTKSDTVQVVDHGDGTGTHVYIAEKDEM 2239
Qy 185 NGTLTCVATWKSILKARKSATVNLT-----VIRCPQD 215
Db 2240 SGTLTAKAKAAGBCETSARVTVNGNKKPBFVQAPQN 2277
RESULT 9
PTK7 HUMAN
ID PTK7_HUMAN STANDARD; PRT; 1070 AA.
AC Q13308; Q13417;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase-like 7 precursor (Colon carcinoma kinase-4)

```

(CCK-4).  
GN PTK7 OR CCK4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Colon carcinoma, and Placenta;  
RC MEDLINE=96074849; PubMed=7478540;  
RX Mossie K., Jallat B., Alves F., Sures I., Plowman G.D., Ullrich A.;  
RA "Colon carcinoma kinase-4 defines a new subclass of the receptor  
RT tyrosine kinase family.";  
RT Oncogene 11:2179-2184(1995).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Fibroblast;  
RC MEDLINE=97037064; PubMed=8882711;  
RX Park S.-K., Lee H.-S., Lee S.-T.;  
RA "Characterization of the human full-length PTK7 cDNA encoding a  
RT receptor protein tyrosine kinase-like molecule closely related to  
RT chick KLG.";  
RL J. Biochem. 119:235-239(1996).  
RN [3]  
RN SEQUENCE FROM N.A., AND REVISION TO 834.  
RA Lee S.-T., Park S.-K., Lee H.-S., Ji A.R., Jung J.W.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY  
CC THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE  
CC PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR  
CC PROGRESSION MARKER.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, LIVER, PANCREAS,  
CC KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID  
CC GLAND, OVARY, BRAIN, HEART AND SKELETAL MUSCLE. ALSO EXPRESSED IN  
CC ERYTHROLEUKEMIA CELLS. BUT NOT EXPRESSED IN COLON.  
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN  
CC RECEPTOR SUBFAMILY.  
CC -!- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U33635; AAA87565.1; -;  
DR EMBL; U40271; AAC50484.2; -;  
DR EMBL; AF447176; AAL39062.1; -;  
DR EMBL; AF447157; AAL39062.1; JOINED.  
DR EMBL; AF447158; AAL39062.1; JOINED.  
DR EMBL; AF447162; AAL39062.1; JOINED.  
DR EMBL; AF447164; AAL39062.1; JOINED.  
DR EMBL; AF447167; AAL39062.1; JOINED.  
DR EMBL; AF447170; AAL39062.1; JOINED.  
DR EMBL; AF447171; AAL39062.1; JOINED.  
DR EMBL; AF447173; AAL39062.1; JOINED.  
DR EMBL; AF447174; AAL39062.1; JOINED.  
DR EMBL; AF447175; AAL39062.1; JOINED.  
DR HSSP; P08631; 1AD5.  
DR Genew; HGNC:9618; PTK7.  
DR MW; 601890; -;  
DR GO; GO:0005886; C:plasma membrane; TAS.  
DR GO; GO:0005888; C:proteoglycan integral to plasma membrane; TAS.  
DR GO; GO:0004714; P:transmembrane receptor protein tyrosine kin. .; TAS.  
DR GO; GO:0007165; P:signal transduction; TAS.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003598; IG c2.  
DR InterPro; IPR003006; IG MHC.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002011; RTKinaseII.

DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00047; ig; 7.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00408; IGC2; 5.  
DR SMART; SM00219; TyrKC; 1.  
DR PROSITE; PS00835; IG-LIKE; 7.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE; PS00239; RECEPTOR TYR\_KIN\_II; FALSE NEG.  
KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;  
KW Immunoglobulin domain; Repeat.  
FT SIGNAL 1 30 POTENTIAL.  
FT CHAIN 31 1070 TYROSINE-PROTEIN KINASE-LIKE 7.  
FT DOMAIN 31 704 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 705 725 POTENTIAL.  
FT DOMAIN 726 1070 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 31 120 IG-LIKE C2-TYPE 1.  
FT DOMAIN 128 218 IG-LIKE C2-TYPE 2.  
FT DOMAIN 225 317 IG-LIKE C2-TYPE 3.  
FT DOMAIN 309 407 IG-LIKE C2-TYPE 4.  
FT DOMAIN 412 497 IG-LIKE C2-TYPE 5.  
FT DOMAIN 503 586 IG-LIKE C2-TYPE 6.  
FT DOMAIN 578 680 IG-LIKE C2-TYPE 7.  
FT DOMAIN 796 1066 PROTEIN KINASE; INACTIVE.  
FT DISULFID 53 101 BY SIMILARITY.  
FT DISULFID 150 200 BY SIMILARITY.  
FT DISULFID 246 301 BY SIMILARITY.  
FT DISULFID 343 391 BY SIMILARITY.  
FT DISULFID 433 481 BY SIMILARITY.  
FT DISULFID 524 570 BY SIMILARITY.  
FT DISULFID 616 664 BY SIMILARITY.  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 92 92 P -> R (IN REF. 2 AND 3).  
FT CONFLICT 147 147 K -> T (IN REF. 2 AND 3).  
FT CONFLICT 207 207 S -> G (IN REF. 2 AND 3).  
FT CONFLICT 495 495 VL -> RV (IN REF. 2 AND 3).  
FT CONFLICT 515 515 G -> E (IN REF. 2 AND 3).  
FT CONFLICT 881 881 E -> G (IN REF. 2 AND 3).  
FT CONFLICT 969 969 A -> P (IN REF. 2 AND 3).  
FT CONFLICT 992 992 S -> F (IN REF. 2 AND 3).  
SQ SEQUENCE 1070 AA; 118260 MW; 47CDF25B9E3698A5 CRC64;  
Query Match 6.0%; Score 124; DB 1; Length 1070;  
Best Local Similarity 25.3%; Pred. No. 0.045;  
Matches 58; Conservative 32; Mismatches 103; Indels 36; Gaps 10;  
Qy 16 VIEGPQNAVTLKGSQARFNCTVS-QGWKLIMWALSDMVLVSRMPEPIINDRFTSQRYD 74  
Db 227 VVLAPQDVVVARYEAMFHCQFSAQPPPSLQWLFEDETPTITNRSPPHLRRATVFA---- 282  
Qy 75 QGGNFTSEMIHNHVPDSDSNIRCSLQNSR-----LHGSAYLTVQVNGELFIPSVNLVVA 129  
Db 283 -----NGSLLLTVQVRPNAGIYRCIGQGQGPPILEATLHLAEIDMPLFEPRVFTAGS 337  
Qy 130 ENEPCEVTCLPSHWTRLPDISWE-LGL-LVSHSSYYVPEPSPDIQSASVILALTPQSN-G 186  
Db 338 EE---RVTLCLPPKGLPEPSVWWEHAGVRLPTHGEVY-----QKGHELVIANTAESDAG 387  
Qy 187 TLTCVATPKSLKARKSATVNLTV-----IRCPQDTGGGINIPGVLSL 229  
Db 388 VYTCHAA--NLAQRRQDVNITVATVPSWLKPKQDSQLEBKGPGYLDCL 434

```
RESULT 10
NTRI HUMAN
ID NTRI HUMAN STANDARD; PRT; 344 AA.
AC Q9P121;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurotrophin precursor (hnt).
GN NT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li G., Jin J., Tan X., Hu S., Yuan J., Qiang B.;
RT "Cloning and identification of human neurotrophin full length cDNA.";
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Neural cell adhesion molecule.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF126426; AF37591.1; -
CC GO; GO:0008038; P:neuronal cell recognition; TAS.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_C2.
CC InterPro; IPR003006; IG_WHC.
CC Pfam; PF00047; IG_3.
CC SMART; SM00408; IgC2; 2.
CC PROSITE; PS50835; IG_LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.
FT SIGNAL 1 31
FT CHAIN 32 344
FT PROPEP 322 344
FT DOMAIN 39 126
FT DOMAIN 136 218
FT DOMAIN 222 309
FT DISULFID 57 115
FT DISULFID 157 201
FT DISULFID 243 295
FT CARBOHYD 44 44
FT CARBOHYD 70 70
FT CARBOHYD 152 152
FT CARBOHYD 284 284
FT CARBOHYD 292 292
FT CARBOHYD 305 305
FT CARBOHYD 321 321
FT LIPID 321 321
FT SEQUENCE 344 AA; 37971 MW; DA4D12C295ABBE3A CRC64;

Query Match 6.0%; Score 123.5; DB 1; Length 344;
Best Local Similarity 23.9%; Pred. No. 0.012;
Matches 63; Conservative 43; Mismatches 105; Indels 53; Gaps 14;

QY 4 VIFLHSGG---SGNEVI-EGPONATVLKGSQARFNCTVSGQKLIWALSDMVVLS--- 55
Db 21 LLLFLVPTGVFVRSQATFPKANDVTVRGESATLCTIDNRTVAVLNRSTILYAGND 80
QY 56 ---VRPMEPIITNDRFTSQDQGNFTGEMIHNVPSDGNIRCSLQ-----NSRLH 106
Db 81 KWLCLPRVLLSN---TQYQY-----SIIQNVVDVDEGPTCSVTQDNHPTKSRVH 129

RESULT 11
NTRI MOUSE
ID NTRI MOUSE STANDARD; PRT; 344 AA.
AC Q99F70;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurotrophin precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim T.H., Choi S.C., Kim J., Jeon J.W., Kim K.D., Lee S.H.;
RT "Cloning and expression of mouse neurotrophin gene in the developing
RT nervous system.";
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Young A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Neural cell adhesion molecule.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF282980; AAK00276.1; -
CC EMBL; BC023307; AAH23307.1; -
```

DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00409; IG; 3.  
DR SMART; SM00408; IGC2; 3.  
DR PROSITE; PS50835; IG LIKE; 3.  
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
KW Repeat; Signal.  
FT SIGNAL 1 31  
FT CHAIN 32 321  
FT PROPEP 322 344  
FT DOMAIN 39 126  
FT DOMAIN 136 218  
FT DOMAIN 222 309  
FT DOMAIN 57 115  
FT DISULFID 157 201  
FT DISULFID 243 295  
FT CARBOHYD 44 44  
FT CARBOHYD 70 70  
FT CARBOHYD 152 152  
FT CARBOHYD 284 284  
FT CARBOHYD 292 292  
FT CARBOHYD 305 305  
FT CARBOHYD 321 321  
FT CARBOHYD 321 321  
FT LIPID 321 321  
FT CONFLICT 75 75  
FT CONFLICT 92 92  
FT CONFLICT 119 119  
FT CONFLICT 187 187  
FT CONFLICT 213 213  
FT CONFLICT 225 225  
SQ SEQUENCE 344 AA; 37984 MW; C885BBA52C148554 CRC64;

Query Match 6.0%; Score 123.5; DB 1; Length 344;

Best Local Similarity 23.9%; Pred. No. 0.012;  
Matches 63; Conservative 43; Mismatches 105; Indels 53; Gaps 14;

Qy 4 VIFLHSGS-----SGNEVI-EGPQATVLKGSQARFNCTVSGWKLIMWALSDMVVL--- 55  
Db 21 LLFLVPTGVPVRSGDATFPKAMDNVTVRQGESATLRCITIDNRTVTRVAVLNRESTILYAGND 80  
Qy 56 ---VRPMEPIITNDRFTSQRYDQGNFTSEMIHNVEPSDSGNIRCSLQ-----NGLRL 106  
Db 81 KWCLDPRVLLSN---TOTQY-----STEQNVVDYDEGPTVCSVQTDNHPKTSRV 129  
Qy 107 GSALVTQVMGELFTPSVNLVAENPECEVCLPSHWTRLPDISWELGLLVSHSXYFVP 166  
Db 130 ---LIVQVSPKIVEISSDISINEGNISLTCIATGRPE-FTVTWR---HSPKAVGFVS 181  
Qy 167 EPSDLQSAVSIALTTPQSNGLTCTVATWKSILKARKSATVNLTVIRCP-----QDTGGGIN 221  
Db 182 EDEYLE---IQGITRQSGEVECSAS-NDVAAPVRRVKVTVNVYPPISEAKGTGVPVG 236  
Qy 222 IPGVL-----SLPSLGFSLPTWKG 241  
Db 237 QKGTLOCEASAVPSAEFQ---WFK 257

RESULT 12

ID OPMC\_HUMAN STANDARD; PRT; 345 AA.  
AC Q14982;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Opioid binding protein/cell adhesion molecule precursor (OBCAM)  
DN (Opioid-binding cell adhesion molecule) (OPCML).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Query Match 5.8%; Score 120; DB 1; Length 345;  
Best Local Similarity 25.0%; Pred. No. 0.023;  
Matches 71; Conservative 37; Mismatches 102; Indels 74; Gaps 18;

Qy 4 VIFLHSGS-----SGNEVI-EGPQATVLKGSQARFNCTVSGWKLIMWALSDMVVL--- 54  
Db 21 LLFLVPTGVPVRSGDATFPKAMDNVTVRQGESATLRCITIDNRTVTRVAVLNRESTILYAGN 79  
Qy 55 ---SVRPMEPIITNDRFTSQRYDQGNFTSEMIHNVEPSDSGNIRCSLQ-----NGLRL 105  
Db 80 DKWSIDPRVILVN---TPTQY-----SIMIQNVVDYDEGPTVCSVQTDNHPKTSRV 128  
Qy 106 HGSAYLTVQVMGELFIPSVNLVAENPECEVTCL-----PSHWTRLPDISWELGLLVSHSS 161

OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Occipital cortex;  
RX MEDLINE=95237612; PubMed=7721093;  
RA Shark K.B., Lee N.M.;  
RT "Cloning, sequencing and localization to chromosome 11 of a cDNA  
RL encoding a human opioid-binding cell adhesion molecule (OBCAM).";  
RL Gene 155:213-217 (1995).  
CC -!- FUNCTION: Binds opioids in the presence of acidic lipids; probably  
CC involved in cell contact.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By  
CC similarity).  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L34774; AAA36387.1; -.  
DR PIR; JC4025; JC4025.  
DR Genew; HGNC:8143; OPCML.  
DR MIM; 600632; -.  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0004985; F:opioid receptor activity; TAS.  
DR GO; GO:0007155; P:cell adhesion; TAS.  
DR GO; GO:0008038; P:neuronal cell recognition; TAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00408; IGC2; 2.  
DR PROSITE; PS50835; IG LIKE; 3.  
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;  
KW Repeat; Signal.  
FT SIGNAL 1 27  
FT CHAIN 28 322

BY SIMILARITY.  
OPIOID BINDING PROTEIN/CELL ADHESION  
MOLECULE.

PROPEP 323 345  
FT DOMAIN 39 126  
FT DOMAIN 136 219  
FT DOMAIN 223 310  
FT DISULFID 57 115  
FT DISULFID 157 202  
FT DISULFID 244 296  
FT CARBOHYD 44 44  
FT CARBOHYD 70 70  
FT CARBOHYD 140 140  
FT CARBOHYD 285 285  
FT CARBOHYD 293 293  
FT CARBOHYD 306 306  
FT LIPID 322 322  
SQ SEQUENCE 345 AA; 38007 MW; E7AD17BEA1AA3FF4 CRC64;

Query Match 5.8%; Score 120; DB 1; Length 345;

Best Local Similarity 25.0%; Pred. No. 0.023;  
Matches 71; Conservative 37; Mismatches 102; Indels 74; Gaps 18;

```

Db 129 H-----LIVQPPQIMNISDTVNEGSSVTLCLLAIGRPE-----PTVTWR-----HLS 173
Qy 162 YY-----FVPEPSDLOSASVIALTPQSGNLTLCVATWKSLSKARKSATVNLTV-----IRC 212
Db 174 VKEQGFEVSEDEYLE-----ISDIKRDQSGEYCSAL-NDVAAPDVRVKITVNPYPYISK 228
Qy 213 PQDTGGGINIPGVLT-----SSLPSLGFLSLPTWKGK-----VGLGLAG 248
Db 229 AKNTGVSQVGGKGLSCASAVPMAEFO---WFKEDTRLATGLDG 269

RESULT 13
ID OPMC BOVIN STANDARD; PRT; 345 AA.
AC P11834;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OB CAM)
DE (Opioid-binding cell adhesion molecule) (OPCML).
GN OPCML OR OB CAM OR OCAM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=99251576; PubMed=2721489;
RA Schofield P.R., McFarland K.C., Hayflick J.S., Wilcox J.N.,
RA Cho T.M., Roy S., Lee N.M., Loh H.H., Seeburg P.H.;
RT "Molecular characterization of a new immunoglobulin superfamily
RL protein with potential roles in opioid binding and cell contact.";
EMBO J. 8:489-495(1989).
CC -I- FUNCTION: Binds opioids in the presence of acidic lipids; probably
CC involved in cell contact.
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -I- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X12672; CAA31192.1; -.
CC PIR; S03199; S03199.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_c2.
CC InterPro; IPR003006; IG_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00408; IGC2; 2.
CC PROSITE; PS50835; IG LIKE; 3.
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 322
FT
FT OPIOID BINDING PROTEIN/CELL ADHESION
FT MOLECULE.
FT PROPEP 323 345
FT DOMAIN 39 126
FT IG-LIKE C2-TYPE 1.
FT DOMAIN 136 219
FT IG-LIKE C2-TYPE 2.
FT DOMAIN 223 310
FT IG-LIKE C2-TYPE 3.
FT POTENTIAL.
FT DISULFID 57 115
FT DISULFID 157 202
FT DISULFID 244 296
FT POTENTIAL.
FT CARBOHYD 44 44
FT N-LINKED (GLCNAC. . .) (PROBABLE).

```

```

FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 293 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 GPI-ANCHOR (POTENTIAL).
SQ SEQUENCE 345 AA; 37914 MW; DIECC8D9E7D8CB19 CRC64;

Query Match 5.8%; Score 119; DB 1; Length 345;
Best Local Similarity 25.1%; Pred. No. 0.027;
Matches 70; Conservative 36; Mismatches 103; Indels 70; Gaps 17;

Qy 4 VIFLHSG-----SGNEVI-EGPQNAVTLKSGQARNFCTVSGWKLIMWALSDMVVL-----54
Db 21 LLFLVPTGVPVRSGDATFPKAMDNVTVRQGESATRLCTIDDRVTRVAV-LNRSTILYAGN 79
Qy 55 ---SVREMEPIITNDRTSORYDQGNFTSEMIHNVPSDSGNIRCSLO-----NSRL 105
Db 80 DKWSIDPRVILVN---TPTQY-----SIMQNVVDYDEGPTCSVDNHPKTSRV 128
Qy 106 HGSAYLTVQVMGELFIPSVNLVVAENEFCEVTCL---PSHWTRLPDISWELGLLVSHSS 161
Db 129 H-----LIVQPPQIMNISDVTVNEGSSVTLCLLAIGRPE-----PTVTWR-----HLS 173
Qy 162 YY-----FVPEPSDLOSASVIALTPQSGNLTLCVATWKSLSKARKSATVNLTV-----IRC 212
Db 174 VKEQGFEVSEDEYLE-----ISDIKRDQSGEYCSAL-NDVAAPDVRVKITVNPYPYISK 228
Qy 213 PQDTGGGINIPGVLT-----SSLPSLGFLSLPTWKGKGLGLA 247
Db 229 AKNTGVSQVGGKGLSCASAVPMAEFO---WFKEDTRLA 264

RESULT 14
ID OPMC RAT STANDARD; PRT; 345 AA.
AC P32736; P32735; Q01653; Q01654;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OB CAM)
DE (Opioid-binding cell adhesion molecule) (OPCML).
GN OPCML OR OB CAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92347701; PubMed=1339369;
RA Lippman D.A., Lee N.M., Loh H.H.;
RA "Opioid-binding cell adhesion molecule (OB CAM)-related clones from a
RA rat brain cDNA library.";
RL Gene 117:249-254(1992).
RN [2]
RP SEQUENCE OF 195-214, AND GPI-ANCHOR.
RX MEDLINE=95198094; PubMed=7891157;
RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,
RA Salzer J.L.;
RA "Cloning of neurotrophin defines a new subfamily of differentially
RA expressed neural cell adhesion molecules.";
RL J. Neurosci. 15:2141-2156(1995).
CC -I- FUNCTION: Binds opioids in the presence of acidic lipids; probably
CC involved in cell contact.
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P32736-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P32736-2; Sequence=VSP_002612;

```

```
CC Note-No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M88710; AAA40859.1; -
CC EMBL; M88711; AAA40860.1; -
CC EMBL; M88709; AAA40858.1; -
CC PIR; JCI238; JCI238.
CC PIR; JCI239; JCI239.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00408; IGC2; 2.
CC PROSITE; PS00835; IG_LIKE; 3.
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
CC Repeat; Signal; Alternative splicing.
CC SIGNAL 1 27
CC CHAIN 28 322
CC MOLECULE
CC PROPEP 323 345
CC DOMAIN 39 126
CC DOMAIN 136 219
CC DOMAIN 223 310
CC DISULFID 57 115
CC DISULFID 157 202
CC DISULFID 244 296
CC CARBOHYD 44 44
CC CARBOHYD 70 70
CC CARBOHYD 140 140
CC CARBOHYD 285 285
CC CARBOHYD 293 293
CC CARBOHYD 306 306
CC LIPID 322 322
CC VARSPLIC 1 27
CC MGVCGLYFLPMKCLVAVSLRLFLVPT -> MYHPAYWVF
CC SATTALLFIP (in isoform 2).
CC /FTid=VSP.002612.
CC SEQUENCE 345 AA; 38067 MW; A3181B0753F95658E CRC64;
Query Match 5.7%; Score 118; DB 1; Length 345;
Best Local Similarity 24.7%; Pred. No. 0.033;
Matches 71; Conservative 38; Mismatches 104; Indels 74; Gaps 18;
QY 4 VIFLHGGG---SGNEVI-EGPONATVLKGSQARFNCTVSGQWKLIMWLSDMVL--- 54
Db 21 LLEFLVPTGVPRSGDATFPAMDVTVRQGESATLRTCTIDRVTVRVAV-INRSTILYAGN 79
QY 55 ---SVREPEPIITNDRTFSQRYDQGNFTSEMIIHNVPESDSGNIRCSLQ-----NSRL 105
Db 80 DKWSIDPRVAILVN---TPTQY-----SIMIONDVYDEGPTCSVQTDNHPKTSRV 128
QY 106 HGSAYLTQVNGELFIFSNVLNVAENPEVTCI-----PSHWTLRDLISWELGLLVSHSS 161
Db 129 H-----LIVQVPPQIMNTSSITVNEISSVTLCLLAIGRPE-----PTVTR-----HLS 173
QY 162 YY-----FVPEPSPDIQSAVSILALTPQSGNTLTVCATWKSILKARKSATVNTIV-----IRC 212
Db 174 VKEGQGVSEDEYLE-----ISDIKEDQSGEYECAL-NDVAAPPVRKVIIVYPPYISK 228
QY 213 PQDTGGGINIPGVL-----SSLPSIGFSLPTWGK-----VGLGLAGTML 251
Db 229 AKNTIGSVGQKGLSCASAVPMAEFQ---WFKEDTRLATGLDGVR 272
```

```
RESULT 15
NEOL_CHICK
ID NEOL_CHICK STANDARD; PRT; 1443 AA.
AC Q90610;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neogenin (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Embryonic brain;
RX MEDLINE=95105243; PubMed=7806578;
RA Vielmetter J., Roman J.M., Dreyer W.J.;
RT "Neogenin, an avian cell surface protein expressed during terminal
RT neuronal differentiation, is closely related to the human tumor
RT suppressor molecule deleted in colorectal cancer.";
RL J. Cell Biol. 127:2009-2020(1994).
CC -!- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DEVELOPMENTAL STAGE: IN RETINA, EXPRESSED ON GANGLION CELL FIBERS
CC AS SOON AS THEY BEGIN TO EXTEND THEIR AXONS.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 6 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U07644; AAC59662.1; -
CC PIR; I50600; I50600.
CC HSPP; P11276; 2MFN.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR003962; FNIII subd.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC Pfam; PF00041; fn3; 6.
CC Pfam; PF00047; Ig; 4.
CC PRINTS; PR00014; FNTYPEIII.
CC SMART; SM00060; FN3; 6.
CC SMART; SM00408; IGC2; 2.
CC PROSITE; PS00835; IG_LIKE; 4.
CC Cell adhesion; Repeat; Transmembrane; Immunoglobulin domain;
CC Glycoprotein.
CC NON TER 1 1
CC DOMAIN <1 1090 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1091 1111 POTENTIAL.
CC DOMAIN 1112 1443 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 18 113 IG-LIKE C2-TYPE 1.
CC DOMAIN 118 204 IG-LIKE C2-TYPE 2.
CC DOMAIN 212 302 IG-LIKE C2-TYPE 3.
CC DOMAIN 307 392 IG-LIKE C2-TYPE 4.
CC DOMAIN 422 519 FIBRONECTIN TYPE-III 1.
CC DOMAIN 522 615 FIBRONECTIN TYPE-III 2.
CC DOMAIN 616 714 FIBRONECTIN TYPE-III 3.
CC DOMAIN 720 814 FIBRONECTIN TYPE-III 4.
CC DOMAIN 835 935 FIBRONECTIN TYPE-III 5.
CC DOMAIN 936 1037 FIBRONECTIN TYPE-III 6.
CC DISULFID 40 95 BY SIMILARITY.
```

FT	DISULFID	139	187	BY SIMILARITY.
FT	DISULFID	236	286	BY SIMILARITY.
FT	DISULFID	328	376	BY SIMILARITY.
FT	CARBOHYD	39	39	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	456	456	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	475	475	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	625	625	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	700	700	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	894	894	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1443 AA;	158050 MW; 558C6795579C0E26 CRC64;	
Query Match 5.7%; Score 117.5; DB 1; Length 1443;				
Best Local Similarity 26.2%; Pred. No. 0.21;				
Matches 49; Conservative 28; Mismatches 73; Indels 37; Gaps 11;				
QY	17	IEGPONATVLKGSQARENCTVSQGW--KLIMWALSDMVVLSVRPMEPIITNDRFTSQRY- 73		
Db	218	VRQPSLLTKVTGQNAVFFC-VAGGFPTFYVRWKNG-----EELITED---SERPA 264		
QY	74	-DQGNFTSEMIHNVEPSDSGNIRC--SLONSRLHGSAYLTVQVMGELFIPSVNLVVAE 130		
Db	265	LRAGGS----LLISDVTEEDVGTYTCIADNENETIEAQAEALAVQVPPPEFLKRPAIYAHE 320		
QY	131	NE----PCEVTCLPSSHWRLPDISW-ELGLLVSHSSYYFVPEPSDLOSASVILALTPOSN 185		
Db	321	SMDIVFECEVTGKPI-----FTVKWKNGDVVIPSDYFKIVKEHNLQ----VGLVKSDE 371		
QY	186	GTLTCVA 192		
Db	372	GFYQCIA 378		

Search completed: November 13, 2003, 03:31:44  
Job time : 12.2686 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 23:19:44 ; Search time 33.7643 Seconds  
(without alignments)  
1795.789 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059

Sequence: 1 MGILVFLHSGSGNEVIEGP.....HPQAFNLASPEKVSNTTVV 382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2059	100.0	382	23 AAU75540	Human B7-like prot
2	2014	97.8	386	23 AAU75542	Human B7-like prot
3	2004	97.3	386	23 AAU75541	Human B7-like prot
4	1514	73.5	377	23 AAU75543	Human B7-like prot
5	1326	64.4	463	22 ABG28169	Novel human diagno
6	924.5	44.9	370	23 AAU75544	Mouse B7-like prot
7	583.5	28.3	631	23 AAU75547	Rat B7-like protei
8	575.5	28.0	270	23 AAU75545	Mouse B7-like prot
9	494	24.0	223	23 AAU75546	Mouse B7-like prot

10	151	7.3	420	23	ABP65012	Human protein SEQ
11	147.5	7.2	404	22	ABP65011	Extracorporeal cir
12	147.5	7.2	404	23	ABP65011	Human protein SEQ
13	147.5	7.2	404	23	AAE23219	Human receptor for
14	147.5	7.2	404	23	AAU77543	Human receptor for
15	147.5	7.2	404	23	AAU77543	Human RAGE protein
16	140	6.8	633	24	AAO26267	gp354 clone protei
17	138	6.7	592	23	ABP75751	Human immunoglobul
18	136	6.7	592	24	AAO26262	Human gp354 protei
19	136	6.6	594	23	ABP75753	Human pancreas GP3
20	136	6.6	594	24	AAO26266	Pancreatic gp354 p
21	136	6.6	708	22	AAU29315	Human PRO polypept
22	136	6.6	708	23	ABP75752	Human gp354 (putat
23	136	6.6	708	24	ABU71403	Human PRO19646 pro
24	136	6.6	708	24	ABU65860	Human secreted/tra
25	136	6.6	708	24	ABU66193	Novel human secret
26	136	6.6	708	24	ABU67697	Human secreted/tra
27	136	6.6	708	24	AAO26285	Human PanCAM prote
28	136	6.6	708	24	ABU65555	Human PRO polypept
29	136	6.6	708	24	ABU58691	Human PRO polypept
30	136	6.6	708	24	ABU56227	Human secreted/tra
31	136	6.6	708	24	ABU57222	Human PRO polypept
32	136	6.6	708	24	ABU10801	Human secreted/tra
33	135	6.6	1496	20	AAW81030	Melanoma associate
34	135	6.6	1496	21	AAU70469	Human p53 target m
35	135	6.6	1496	24	ABU03498	Angiogenesis-assoc
36	135	6.6	1498	22	ABP11587	Human peroxidase
37	130	6.3	4675	23	ABP70085	Human NOV13b. Hom
38	130	6.3	4691	23	ABP70084	Human NOV13a. Hom
39	128	6.2	1447	16	AAW68553	Deleted in colorec
40	128	6.2	1447	20	AAU33498	Human DCC protein.
41	128	6.2	1447	22	AAU50693	Human UNC-40 prote
42	128	6.2	1447	24	ABU19765	Human MP21 protein
43	128	6.2	1728	12	AAU13144	Deleted in Colorec
44	128	6.2	1953	23	AAU84351	Protein MYLK diffe
45	127.5	6.2	467	19	AAW62575	Alternatively apli

#### ALIGNMENTS

RESULT 1  
AAU75540  
L AAU75540 standard; Protein; 382 AA.

AC AAU75540;

DT 23-APR-2002 (first entry)

DE Human B7-like protein, B7-L\_hl.

Human; B7-like protein; B7-L; antiinfertility; gynaecological;  
antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
antiinflammatory; dermatological; antiporiatic; antiprotective;  
antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
antiaesthetic; nephrotropic; antibacterial; virucide; tumour; cancer;  
reproductive disorder; graft versus host disease; autoimmune disease;  
toxic shock syndrome; allergy; nephropathy; skin disorder;  
endocrinopathy; lymphoproliferative disorder.

OS Homo sapiens.

FN WO200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.

XX 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.



CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allo sensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, colliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the amino acid sequence of human B7-L<sub>h3</sub>.  
 XX  
 SQ Sequence 386 AA;  
 Query Match 97.8%; Score 2014; DB 23; Length 386;  
 Best Local Similarity 99.7%; Pred. No. 2.3e-159;  
 Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 GSGSGNEVIEGPQNAVTLKGSQARFNCVTSQGWKLIMWALSDMVLSVRPMEPIITNDRF 68  
 DB 13 GSGSGNEVIEGPQNAVTLKGSQARFNCVTSQGWKLIMWALSDMVLSVRPMEPIITNDRF 72  
 QY 69 TSQRYDQGNFTSEMIHNVEPDSGNIRCSLQNSRLHGSAYLTVQVWGELFIPSVNLV 128  
 DB 73 TSQRYDQGNFTSEMIHNVEPDSGNIRCSLQNSRLHGSAYLTVQVWGELFIPSVNLV 132  
 QY 129 AENPECEVTCLPSHWTRLPTDLSWELGLLVSHSSYFVPEPSDLSQSAVILALTQSNQTL 188  
 DB 133 AENPECEVTCLPSHWTRLPTDLSWELGLLVSHSSYFVPEPSDLSQSAVILALTQSNQTL 192  
 QY 189 TCVAATWSLARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGSFSLPTWKGVLGLIAG 248  
 DB 193 TCVAATWSLARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLGSFSLPTWKGVLGLIAG 252  
 QY 249 TMLTPTCTLIRCCCRRCRCCGCCNCCRCFCRRRKGFRIFQKKSEKTKETETE 308  
 DB 253 TMLTPTCTLIRCCCRRCRCCGCCNCCRCFCRRRKGFRIFQKKSEKTKETETE 312  
 QY 309 SGNENSGYNSDEQKTTTASLPKSCSSDPQQRNNSCGPPHQADQPPPPASHQAS 368  
 DB 313 SGNENSGYNSDEQKTTTASLPKSCSSDPQQRNNSCGPPHQADQPPPPASHQAS 372  
 QY 369 NLASPEKVSNTTV 382  
 DB 373 NLASPEKVSNTTV 386  
 RESULT 3  
 ID AAU75541 standard; Protein; 386 AA.  
 XX  
 AC AAU75541;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE Human B7-like protein, B7-L<sub>h2</sub>.  
 XX  
 KW Human, B7-like protein; B7-L; antiinfertility; gynaecological;  
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
 KW antiaesthetic; nephrotropic; antibacterial; virucide; tumour; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder.  
 XX  
 OS Homo sapiens.  
 XX

PN W0200200710-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 28-JUN-2001; 2001WO-US20719.  
 XX  
 FR 28-JUN-2000; 2000US-214512P.  
 PR 28-NOV-2000; 2000US-0729264.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
 XX WPI; 2002-130881/17.  
 DR N-PSDB; ABK13029.  
 XX  
 PT New B7-like polypeptides, polynucleotides and their modulators, useful  
 PT for diagnosing, preventing and treating reproductive, immune and  
 PT proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX  
 PS Claim 13; Fig 2; 135pp; English.  
 XX  
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 CC The polypeptide polynucleotide encoding it and antibody against (I) are  
 CC useful for treating B7-like polypeptide-related disease, disorders or  
 CC conditions including reproductive disorders (e.g. infertility, and  
 CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (I) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allo sensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, colliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the amino acid sequence of human B7-L<sub>h2</sub>.  
 XX  
 SQ Sequence 386 AA;  
 Query Match 97.3%; Score 2004; DB 23; Length 386;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-157;  
 Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 9 GSGSGNEVIEGPQNAVTLKGSQARFNCVTSQGWKLIMWALSDMVLSVRPMEPIITNDRF 68  
 DB 13 GSGSGNEVIEGPQNAVTLKGSQARFNCVTSQGWKLIMWALSDMVLSVRPMEPIITNDRF 72  
 QY 69 TSQRYDQGNFTSEMIHNVEPDSGNIRCSLQNSRLHGSAYLTVQVWGELFIPSVNLV 128  
 DB 73 TSQRYDQGNFTSEMIHNVEPDSGNIRCSLQNSRLHGSAYLTVQVWGELFIPSVNLV 132

QY 129 AENEPCVETCLPSHWTRLPDISWELGLLVSHSSYYFVPEPSDLQSAVSIILALTPQSNGL 188  
Db 133 AENEPCVETCLPSHWTRLPDISWELGLLVSHSSYYFVPEPSDLQSAVSIILALTPQSNGL 192  
QY 189 TCVATWKSARKSATVNLTVIRCPDQTGGGINTPGVLSLPSLPGSLPTWGWKVLGLAG 248  
Db 193 TCVATWKSARKSATVNLTVIRCPDQTGGGINTPGVLSLPSLPGSLPTWGWKVLGLAG 252  
QY 249 TMLLTPTCTLTIRCCRRRCGCCGCCRCFCFCRRKRGRFRIQFKKSEKKTNKETETE 308  
Db 253 TMLLTPTCTLTIRCCRRRCGCCGCCRCFCFCRRKRGRFRIQFKKSEKKTNKETETE 312  
QY 309 SGNENSGYNSDEQKTTDTASLPKSCSSDPQQRNSCGPPHQRADQRRPPASHPOASF 369  
Db 313 SGNENSGYNSDEQKTTDTASLPKSCSSDPQQRNSCGPPHQRADQRRPPASHPOASF 372  
QY 369 NLASPEKVSNTTVV 382  
Db 373 NLASPEKVSNTTVV 386

RESULT 4

AAU75543  
ID AAU75543 standard; Protein; 377 AA.

XX AAU75543;

XX 23-APR-2002 (first entry)

XX Human B7-like protein, B7-L\_h4.

XX Human; B7-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;  
KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; graft versus host disease; autoimmune disease;  
KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder.

XX Homo sapiens.

XX W0200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.

XX 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.

XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX WPI; 2002-130881/17.

XX N-PSDB; ABK13031.

XX New B7-like polypeptides, polynucleotides and their modulators, useful  
PT for diagnosing, preventing and treating reproductive, immune and  
PT proliferative disorders, e.g. cancer and arteriosclerosis -

PS Claim 13; Fig 4; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
CC useful for treating B7-like polypeptide-related disease, disorders or  
CC conditions including reproductive disorders (e.g. infertility  
CC miscarriage, preterm labour and delivery and endometriosis) and  
CC proliferative disorders. Antibodies, soluble proteins comprising  
CC extracellular domains and other regulators of B7-L polypeptides are  
CC useful for enhancing the immune response to tumours. (I) plays a role in  
CC growth and maintenance of cancer cells based on the observation of

CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
CC polypeptide. Hence modulators of (I) are useful for the treatment of  
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
CC in allograft transplantation, graft versus host disease, T-cell  
CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
CC molecules are useful for alleviating the symptoms associated with  
CC diseases involving chronic immune cell dysfunction or to treat  
CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
CC purpura and psoriasis, chronic inflammatory disease such as  
CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
CC are also useful as immunosuppressive agents for bone marrow and organ  
CC transplantation or to prolong graft survival. B7-L molecules are also  
CC useful for diagnosis and treatment of diseases involving abnormal cell  
CC proliferation, including arteriosclerosis and vascular restenosis.  
CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
CC shock syndrome or allosensitisation due to blood transfusions, and for  
CC treatment of allergy, asthma and hypersensitivity reactions.  
CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
CC pemphigoid), endocrinopathies (grave's disease), various pneumopathies  
CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
CC lymphoproliferative disorders such as multiple myeloma. The present  
CC sequence represents the amino acid sequence of human B7-L\_h4.

XX SQ Sequence 377 AA;

Query Match 73.5%; Score 1514; DB 23; Length 377;

Best Local Similarity 88.2%; Pred. No. 5.6e-117;

Matches 284; Conservative 12; Mismatches 20; Indels 6; Gaps 2;

QY 9 GSGSGNEVIEGPQNAVTLKSGQARENCTVSGQNKLIIMWALSDMVLSVRPWEPIITNDRF 68  
Db 13 GSGSGNEVIEGPQNAVTLKSGQARENCTVSGQNKLIIMWALSDMVLSVRPWEPIITNDRF 72  
QY 69 TSQRYDQGGNFTSEMIHNVPEPSDGNIRCSLQNSRLHGSAYLTVQVMGELFIPSVNLV 128  
Db 73 TSQRYDQGGNLTSEMIHNVPEPSDGNIRCSLQNSRLHGSAYLTVQVMGELFIPSVNLV 132  
QY 129 AENEPCVETCLPSHWTRLPDISWELGLLVSHSSYYFVPEPSDLQSAVSIILALTPQSNGL 188  
Db 133 AENEPCVETCLPSHWTRLPDISWELGLLVSHSSYYFVPEPSDLQSAVSIILALTPQSNGL 192  
QY 189 TCVATWKSARKSATVNLTVIRCPDQTGGGINTPGVLSLPSLPGSLPTWGWKVLGLAG 248  
Db 193 TCVATWKSARKSATVNLTVIRCPDQTGGGINTPGVLSLPSLPGSLPTWGWKVLGLAG 252  
QY 249 TMLLTPTCTLTIRCCRRRCGCCGCCRCFCFCRRKRGRFRIQFKKSEKKTNKETETE 308  
Db 253 TMLLTPTCTLTIRCCRRRCGCCGCCRCFCFCRRKRGRFRIQFKKSEKKTNKETETE 309  
QY 309 SGNENSGYNSDEQKTTDTASLP 330  
Db 310 T---PATIQMKRPQTPFLSLP 328

RESULT 5

ABG28169

ID ABG28169 standard; Protein; 463 AA.

XX ABG28169;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #28160.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX Homo sapiens.





QY 161 -----SYFVPEPSDIQ 172  
Db 312 LSVRADFPTEITEYGNPSADIKRITCFASGGPKPRLGWLGNRELSNPSLEPCNFM 371  
QY 173 SAVSILALTPQSGTLCVATWKSLSKARKSATVNL----- 207  
Db 372 RVLSQLDLPLGLNGTLTCVAELDLQASKSLTVNLGINTISODPESELYTSSQLDFNA 431  
QY 208 -----TVIRCPQDTGGGINIPGVLSLSPSGSLPTWKGVLGL 246  
Db 432 TYDHFIDCFIEYDAHVSNQFTVQPPD-----SIGEGPALPTWAILLAV 479  
QY 247 AGTMLL-----TPT 255  
Db 480 AFSLLILILVLIIFTWKPPDPDEKQTVPFAMAGPDAVKAIIFIAITVIAVIA 539  
QY 256 CYLTIRCCCRRCGCCNCCRCFCRRKGRFIOFO-KKSEKENTK---ETETESGN 311  
Db 540 IAIIFPCCCAAR-----REKESTYQNEIRKSNMNTNKADPETKLKSGK 595  
QY 312 ENSGVNSDEQKTTDTASLPKSCESDPEQRNSSCGPPHQRADQ 355  
Db 586 ENVGYSDECIIVKPRRCFRRNEASRETWNKLYIGPVBAABQ 629

RESULT 8  
AAU75545  
ID AAU75545 standard; Protein; 270 AA.

AC AAU75545;  
XX  
XX  
XX 23-APR-2002 (first entry)  
DE Mouse B7-like protein, B7-L\_m2.  
XX  
XX Mouse; B7-like protein; B7-L; antiinfertility; gynaecological;  
KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
KW antidiabetic; haemostatic; antithyroid; antidiulcer; antiallergic;  
KW antitasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
KW reproductive disorder; allergy versus host disease; autoimmune disease;  
KW toxic shock syndrome; graft versus host disease; skin disorder;  
KW endocrinopathy; lymphoproliferative disorder.

XX Mus musculus.

XX WO200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.

XX 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.

PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

XX WPI; 2002-130881/17.

DR N-PSDB; ABK13033.

XX New B7-like polypeptides, polynucleotides and their modulators, useful

XX for diagnosing, preventing and treating reproductive, immune and

XX proliferative disorders, e.g. cancer and arteriosclerosis -

XX Claim 13; Fig 6; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).

XX The polypeptide, polynucleotide encoding it and antibody against (I) are

XX useful for treating B7-like polypeptide-related disease, disorders or

XX conditions including reproductive disorders (e.g. infertility,

CC miscarriage, preterm labour and delivery and endometriosis) and  
CC proliferative disorders. Antibodies, soluble proteins comprising  
CC extracellular domains and other regulators of B7-L polypeptides are  
CC useful for enhancing the immune response to tumours. (i) plays a role in  
CC growth and maintenance of cancer cells based on the observation of  
CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
CC polypeptide. Hence modulators of (i) are useful for the treatment of  
CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
CC in allograft transplantation, graft versus host disease, T-cell  
CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
CC molecules are useful for alleviating the symptoms associated with  
CC diseases involving chronic immune cell dysfunction or to treat  
CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
CC purpura and psoriasis, chronic inflammatory disease such as  
CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. they  
CC are also useful as immunosuppressive agents for bone marrow and organ  
CC transplantation or to prolong graft survival. B7-L molecules are also  
CC useful for diagnosis and treatment of diseases involving abnormal cell  
CC proliferation, including arteriosclerosis and vascular restenosis.  
CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
CC shock syndrome or allosensitisation due to blood transfusions, and for  
CC treatment of allergy, asthma and hypersensitivity reactions,  
CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
CC lymphoproliferative disorders such as multiple myeloma. The present  
CC sequence represents the amino acid sequence of mouse B7-L\_m2.

XX Sequence 270 AA;

Query Match 28.0%; Score 575.5; DB 23; Length 270;  
Best Local Similarity 35.1%; Pred. No. 1.9e-39;  
Matches 136; Conservative 42; Mismatches 73; Indels 137; Gaps 6;  
QY 3 LVIF--LHGSGGNEVIEGPQNAVTLKGSQARENCTVSQGWKLIMWALSDMVLVSVPME 60  
Db 12 LVLAQLTASGSSYQIIEGPQNVTLKDEAHFNCTVTHGKLLMTLQNMVLSLTQ 71  
QY 61 PIITNDRFTSQYDQGNFTSEMIHNVPSDSGNTRCSLQNSRLHGSAYLTQVMGELF 120  
Db 72 PIITNRRFTYASYNSTDSTFISELIHVDQVSDSGSVQCSLQNSHGFGSAFLSVQ----- 125  
QY 121 IPSWNLVVAENEPCEVTCLPSHWTLPDISWELGLLVSHSSYVFPSPDLQSAVSLAL 180  
Db 126 ----- 125  
QY 181 TPQSNGLTCAVATWKSLSKARKSATVNLTVIRCPDQTGGGINIPGVLSLSPSLGFLPTWG 240  
Db 126 -----DSIGEFGALPTWA 139  
QY 241 KVGLGLAGTWLTPCTLTIRCCRRRCGCCNCCRCFCRRKGRFIOFO---KKSE 297  
Db 140 IILLAVAFSILLIILVLIIFCC-----CASRRKEESTYQNEIRKSA 184  
QY 298 KEKTNK---ETETESGNNSGNSYNSDEQKTTDTASLPKSCESDPEQRNSSCGPPHQRAD 354  
Db 185 NWRTNKADPETKLKSGKENYGVSSDEAKAAQATSLPDKSAEVLPEKRSSSL---PYQELN 242  
QY 355 QRPPRPASHPQASFNLASPEKVSNTTVV 382  
Db 243 KHQPGATHPRVSVFDIASPQKVRNVLV 270

RESULT 9  
AAU75546  
ID AAU75546 standard; Protein; 223 AA.  
XX  
XX AAU75546;  
AC

XX	23-APR-2002 (first entry)	CC	sequence represents the amino acid sequence of mouse B7-L_m3.
XX		XX	
XX	Mouse B7-like protein, B7-L_m3.	XX	
XX		XX	
XX	Mouse; B7-like protein; B7-L; antiinfertility; gynaecological;	XX	
XX	antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;	XX	
XX	antifungal; dermatological; antipsoriatic; neuroprotective;	XX	
XX	antidiabetic; haemostatic; antithyroid; antitumor; cancer;	XX	
XX	antidiabetic; nephrotropic; antibacterial; virucide; tumour; cancer;	XX	
XX	reproductive disorder; graft versus host disease; autoimmune disease;	XX	
XX	toxic shock syndrome; allergy; nephropathy; skin disorder;	XX	
XX	endocrinopathy; lymphoproliferative disorder.	XX	
XX		XX	
XX	Mus musculus.	XX	
XX		XX	
XX	WO200200710-A2.	XX	
XX		XX	
XX	03-JAN-2002.	XX	
XX		XX	
XX	28-JUN-2001; 2001WO-US20719.	XX	
XX		XX	
XX	28-JUN-2000; 2000US-214512P.	XX	
XX		XX	
XX	28-NOV-2000; 2000US-0729264.	XX	
XX		XX	
XX	(AMGE-) AMGEN INC.	XX	
XX		XX	
XX	Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;	XX	
XX		XX	
XX	WPI; 2002-130881/17.	XX	
XX		XX	
XX	N-PSDB; ABK13034.	XX	
XX		XX	
XX	New B7-like polypeptides, polynucleotides and their modulators, useful	XX	
XX	for diagnosing, preventing and treating reproductive, immune and	XX	
XX	proliferative disorders, e.g. cancer and arteriosclerosis -	XX	
XX		XX	
XX	Claim 13; Fig 7; 135pp; English.	XX	
XX		XX	
XX	The invention relates to an isolated B7-like (B7-L) polypeptide (I).	XX	
XX	The polypeptide, polynucleotide encoding it and antibody against (I) are	XX	
XX	useful for treating B7-like polypeptide-related disease, disorders or	XX	
XX	conditions including reproductive disorders (e.g. infertility,	XX	
XX	miscarriage, preterm labour and delivery and endometriosis) and	XX	
XX	proliferative disorders. Antibodies, soluble proteins comprising	XX	
XX	extracellular domains and other regulators of B7-L polypeptides are	XX	
XX	useful for enhancing the immune response to tumours. (I) plays a role in	XX	
XX	growth and maintenance of cancer cells based on the observation of	XX	
XX	seminal vesicle hyperplasia in transgenic mice overexpressing B7-L	XX	
XX	polypeptide. Hence modulators of (I) are useful for the treatment of	XX	
XX	cancer including seminal vesicle cancer, lung, brain, breast, ovarian,	XX	
XX	testicular cancer and cancers of haematopoietic system. B7-L polypeptide	XX	
XX	pathway can be manipulated to regulate cytotoxic T-lymphocyte response	XX	
XX	in allograft transplantation, graft versus host disease, T-cell	XX	
XX	dependent B-cell mediated diseases and autoimmune diseases. B7-L	XX	
XX	molecules are useful for alleviating the symptoms associated with	XX	
XX	diseases involving chronic immune cell dysfunction or to treat	XX	
XX	autoimmune diseases such as systemic lupus erythematosus, rheumatoid	XX	
XX	arthritis, multiple sclerosis, diabetes, immune thrombocytopenic	XX	
XX	purpura and psoriasis, chronic inflammatory disease such as	XX	
XX	inflammatory bowel disease (Crohn's disease and ulcerative colitis),	XX	
XX	Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They	XX	
XX	are also useful as immunosuppressive agents for bone marrow and organ	XX	
XX	transplantation or to prolong graft survival. B7-L molecules are also	XX	
XX	useful for diagnosis and treatment of diseases involving abnormal cell	XX	
XX	proliferation, including arteriosclerosis and vascular restenosis.	XX	
XX	Antagonists of B7-L polypeptides are useful for alleviation of toxic	XX	
XX	shock syndrome or allo sensitisation due to blood transfusions, and for	XX	
XX	treatment of allergy, asthma and hypersensitivity reactions.	XX	
XX	nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and	XX	
XX	penicilliosis), endocrinopathies (Grave's disease), various pneumopathies	XX	
XX	(extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,	XX	
XX	thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and	XX	
XX	lymphoproliferative disorders such as multiple myeloma. The present	XX	

CC	sequence represents the amino acid sequence of mouse B7-L_m3.
XX	
XX	Sequence 223 AA;
XX	
XX	Query Match 24.0%; Score 494; DB 23; Length 223;
XX	Best Local Similarity 31.9%; Pred. No. 8.4e-33;
XX	Matches 123; Conservative 37; Mismatches 47; Indels 178; Gaps 7;
XX	
QY	3 LVIF--LHSGSGNEVIEGPQNATVLKGSQARFNCTVSQGWKLIMWALSMDVLSVRPME 60
DB	12 LVILAQLTASGSSYQIIIEGPQNTVLKDSAEHFNCTVTHGWKLIMWLTQMVLSTTQ 71
QY	61 PIITNDRFTSQYDQGNFTSEMIHNVPDSGNRCISLQNSRLHGSAYLTVQWGELEP 120
DB	72 PIITNDRFTSQYDQGNFTSEMIHNVPDSGNRCISLQNSRLHGSAYLTVQWGELEP 125
QY	121 IFSVNLVVAENEPCEVTCLPSHWTRLPDISWELGLLVSHSYVFPVPEPSDLQSAVSIAL 180
DB	126 -----ESTY-----QNEI----- 133
QY	181 TPQNGTLTCVATWKSARKKATVNLTVIRCPQDTGGGINTPGVLSLPSLGLSPTWG 240
DB	134 -----RKSANM----- 139
QY	241 KVGLGLAGTWLTPCTCTLTIRCCCRRCGCCGCCFCCRRKRGFRIOFKKSEK 300
DB	140 ----- 140
QY	301 TNK---ETETSGNENSGYNSDEQKTTDTASLPKSCSSDPPQNRSSCGPPHQRADQRP 357
DB	141 TNKADPTKLSGKNGYSSDEAKAAQTASLPKSAEVSLEPKRSSSL--PYQELNKHQ 198
QY	358 PRPASHPQASFNLASPEKVSNTTV 382
DB	199 PGPATHPRVSPDIASPKVRNLT 223
XX	
XX	RESULT 10
XX	ABP65012
XX	ID ABP65012 standard; Protein; 420 AA.
XX	AC ABP65012;
XX	DT 28-MAR-2003 (first entry)
XX	DE Human protein SEQ ID 672.
XX	Human; expressed sequence tag; EST;
XX	haematopoietic disorder; central nervous system disease; viral infection;
XX	peripheral nervous system disease; non-healing wound; infectious disease;
XX	immune deficiency; immune disorder; bacterial infection; allergy; cancer;
XX	antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
XX	cytostatic; haemostatic; virucide; antibacterial; fungicide;
XX	immunostimulant; cerebroprotective.
XX	Homo sapiens.
XX	WO200259260-A2.
XX	01-AUG-2002.
XX	16-NOV-2001; 2001WO-US42950.
XX	17-NOV-2000; 2000US-0714936.
XX	(HYSE-) HYSEQ INC.
XX	Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
XX	Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX	WPI; 2002-590824/63.
XX	N-PSDB; ABQ99598.



Db 390 -QSEEPAGESTGGP 404

RESULT 12

ABP65011 standard; Protein; 404 AA.

XX AC ABP65011;

XX DT 25-FEB-2003 (first entry)

XX DE Human protein SEQ ID 671.

XX KW Human; expressed sequence tag; EST;

XX KW haematopoietic disorder; central nervous system disease; viral infection;

XX KW peripheral nervous system disease; non-healing wound; infectious disease;

XX KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;

XX KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;

XX KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;

XX KW cytostatic; haemostatic; virucide; antibacterial; fungicide;

XX KW immunostimulant; cerebroprotective.

OS Homo sapiens.

XX WO200259260-A2.

XX PN 01-AUG-2002.

XX PD 16-NOV-2001; 2001WO-US42950.

XX PF 17-NOV-2000; 2000US-0714936.

XX PR (HYSE-) HYSEQ INC.

XX PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;

XX PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX DR WPI; 2002-530824/63.

XX DR N-PSDB; ABQ99597.

XX PT New isolated polynucleotide, useful in research, diagnostic or

XX PT therapeutic methods, e.g. preventing or treating disorders involving

XX PT aberrant protein expression or biological activity -

XX PS Claim 20; SEQ ID 671; 394pp; English.

XX CC The present invention relates to novel human coding sequences

XX CC (ABQ99268-ABQ99508) and proteins (ABP64682-ABP65022). The sequences are

XX CC useful in therapeutic, diagnostic and research methods. The

XX CC polynucleotides may be used in the field of molecular biology as

XX CC hybridisation probes, primers for PCR, for chromosome and gene mapping,

XX CC for the recombinant production of protein, or in generation of anti-sense

XX CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed

XX CC sequence tags (ESTs) for identifying expressed genes or for physical

XX CC mapping of the human genome. The proteins may be used as molecular weight

XX CC markers, or as nutritional sources or supplements. The proteins may be

XX CC used to maintain and expand cell population in a totipotent or

XX CC pluripotent state useful for re-engineering damaged or diseased

XX CC tissues, transplantation, manufacture of bio-pharmaceuticals or the

XX CC development of bio-sensors. The polynucleotides and proteins are useful

XX CC for preventing, treating or ameliorating disorders involving aberrant

XX CC protein expression or biological activity, e.g. haematopoietic disorders,

XX CC central/peripheral nervous system diseases, mechanical and traumatic

XX CC disorders, non-healing wounds, immune deficiencies and disorders,

XX CC infectious diseases caused by viral, bacterial or fungal infection,

XX CC autoimmune disorders, allergic reactions and conditions, coagulation

XX CC disorders, or cancer. The polynucleotide sequences of the invention

XX CC assembled from ESTs isolated mainly by sequencing by hybridisation, and

XX CC in some cases, sequences obtained from one or more public databases.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 404 AA;

Query Match 7.2%; Score 147.5; DB 23; Length 404;

Best Local Similarity 22.6%; Pred No. 0.00088;

Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

QY 15 EVIEGFQNAVTLKGSQARFNCTVSQG---WKLIMWALSDMVLSVRMPFIITNDRFTS- 70

DB 125 EIVDSASELTA--GVNPKVGTGCVSEGSYPAGTILSWHLDG-----KPLVPNEKGVSV 173

QY 71 ----QRYDQGNFT--SEMIHNVFSDSGNIR-----CSLQNSLHGSAYLTQVVMGELF 120

DB 174 KEQTRRHPETGLFTLQSELM---VTARGGDPRPTFCSPGLPRHRLRTAPIQPRVW 230

QY 121 IP-----SVNLVVAENPP-----CEVTCLPSSHWRPLDPDISWELGLIVSHSSY 162

DB 231 EPVPLEEVQLV---EPEGGAVAPGTVTLTCEVPAQS-----PQIHWKMD----- 274

QY 163 YFVPEPSDLQSAVSIILATPQSNGTITCVATWLSLKARKSATVNLTVIRPQDTGGGINI 222

DB 275 -GVPLPLPPSPVLILPEIGPDQGTYSVATHSSHQPSRAVSISII-PEGEG----- 327

QY 223 PGVLSLPSLGFSLPTWGVKVLGLAGTMLLT-----PTCLTIRCCCRRCGCCGCC 275

DB 328 -----PTAGSVGGSLGLTALALGILGLGTAALLIGVILWQR----- 366

QY 276 CRCCFCRRKRGRFRIQFKKSEKKT--NKETETESGNENSGYNSDEQKTTDTASLPKS 333

DB 367 -----QREGEEKAPENQEEBERAELN----- 389

QY 334 CESSDPEQRNSCGPP 349

DB 390 -QSEEPAGESTGGP 404

RESULT 13

AAE23219

ID AAE23219 standard; Protein; 404 AA.

XX AC AAE23219;

XX DT 27-AUG-2002 (first entry)

XX DE Human receptor for advanced glycosylation end product (RAGE) protein.

XX KW Human; Receptor for advanced glycosylation end product; RAGE; cardiact;

XX KW tissue growth; neointimal formation; blood vessel; restenosis; diabetes;

XX KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;

XX KW transgenic animal; acute thrombotic stroke; venous thrombosis.

XX OS Homo sapiens.

XX WO200230889-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US32036.

XX PR 13-OCT-2000; 2000US-0687528.

XX PA (UYCO ) UNIV COLUMBIA NEW YORK.

XX PI Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;

XX DR WPI; 2002-426260/45.

XX DR N-PSDB; AAD36952.

XX PT Inhibiting new tissue growth or neointimal formation in blood vessels

XX PT of subject suffering from diabetes, stroke and preventing restenosis,

XX PT comprises administering inhibitor of receptor for advance glycation end

XX PT product -

XX PS Disclosure; Page 16; 43pp; English.

XX The invention relates to a method for inhibiting new tissue growth or  
CC neointimal formation in blood vessels in a subject that has experienced  
CC blood vessel injury and preventing exaggerated restenosis in a diabetic  
CC subject. The method comprises administering an inhibitor of receptor for  
CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new  
CC tissue growth or neointimal formation in subject's blood vessels and  
CC preventing restenosis in the subject. The method is useful for inhibiting  
CC new tissue growth or neointimal formation in blood vessels in a subject  
CC like non-human animal, a transgenic non-human animal or a human suffering  
CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable  
CC angina, myocardial infarction, abrupt closure following angioplasty or  
CC stent placement, or thrombosis as a result of peripheral vascular surgery  
CC The method is also useful for preventing restenosis and for determining  
CC whether a compound inhibits new tissue growth in a blood vessel in a  
CC subject. The present sequence is human receptor for advanced  
CC glycosylation end product (RAGE) protein.

XX Sequence 404 AA;

Query Match 7.2%; Score 147.5; DB 23; Length 404;  
Best Local Similarity 22.6%; Pred. No. 0.00089;  
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

QY 15 EVIEGPNATVLKGSQARFNCTVSQ---WKLIMWALSDMVLSVRPMEPIITNDRTS- 70  
Db 125 EIVDSASELTA--GVNPKVGTCSGSPAGTLSWHLG-----KELVPEKGVSV 173  
QY 71 ----QRYDQGNFT--SEMIHNVPESDSGNIR-----CSLQNSRLHGSAYLTVQVMGELF 120  
Db 174 KEQTRRHPTGLFTLOSELN---VTPARGDPRPTFCSPGLPRHRLTAPIQPRVW 230  
QY 121 IP-----SVNLVVAENEP-----CEVTCLPSHWTRLPDLSWELGLLVSHSSY 162  
Db 231 EPVPLEEVQLW---EPEGGAVAPGGTVTLTCEVPAQS-----PQHWNKD----- 274  
QY 163 YFVPEPSDLQSAVSILALTPQSNGLTCVATWLSKLKARKSATVNLTVIRCPQDTGGGINI 222  
Db 275 -GVPLPLPSPVLILPEIGPDQGTVCVATHSHGQESRAVSISIIIE-PGEEG----- 327  
QY 223 PGLVSLPSLPSLPTWKGVLGLAGTMLLT-----PTCTLTIRCCCRCCGCGNCC 275  
Db 328 -----PTAGSVGGSGLGTALALGILGLGTAALLIGVILWQRR----- 366  
QY 276 CRCCFCCRRKRGFRIQFKKSEKKT--NKETETESGNSGNSGYNSEDEQKTTDTASLPKPS 333  
Db 367 -----QRRGERKAPENQEEEEERAEALN----- 389  
QY 334 CRSSDPEQRNNSCGPP 349  
Db 390 -QSEEPAGESSTGGP 404

RESULT 14

AAU77543

ID AAU77543 standard; Protein; 404 AA.

XX

XX

AC AAU77543;

XX

XX

05-JUN-2002 (first entry)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PD 21-FEB-2002.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

&lt;

AAW48745  
ID AAW48745 standard; protein; 404 AA.  
AC AAW48745;  
XX  
DT 02-APR-2002 (first entry)  
DE Human RAGE protein SEQ ID NO 1.  
KW Human; RAGE; receptor for advanced glycosylated endproduct; receptor;  
KW antidiabetic; neuroprotective; cytostatic; antiinflammatory; vasotropic;  
KW nephrotropic; dermatological; antiarteriosclerotic; nootropic; diabetes;  
KW Alzheimer's disease; cancer; inflammation; kidney failure;  
KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.  
XX  
OS Homo sapiens.  
XX  
PN WC200192892-A2.  
XX  
PD 06-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US17447.  
XX  
PR 30-MAY-2000; 2000US-207342P.  
PR 05-MAR-2001; 2001US-0799152.  
XX  
PA (TRAN-) TRANS TECH PHARMA.  
XX  
PI Shabbaz M;  
XX  
DR WPI; 2002-114372/15.  
XX  
PT Detecting a receptor for advanced glycosylated endproducts (RAGE)  
PT modulators, for treating e.g., cancer, diabetes or inflammation,  
PT comprises measuring the amount of bound anti-RAGE antibody -  
XX  
PS Claim 1; Fig 2; 49pp; English.  
XX  
CC The invention relates to detecting receptor for advanced glycosylated  
CC endproducts (RAGE) modulators comprises determining the amount of RAGE  
CC protein or its fragment bound to the pre-adsorbed ligand by measuring the  
CC amount of anti-RAGE antibody bound to the solid surface. The method is  
CC useful for rapid, high-throughput identification of compounds that  
CC modulate RAGE. The compounds are useful for treating symptoms of diabetes  
CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's  
CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis  
CC or inflammatory lupus nephritis, erectile dysfunction and  
CC atherosclerosis.  
XX  
SQ Sequence 404 AA;

Query Match 7.2%; Score 147.5; DB 23; Length 404;  
Best Local Similarity 22.6%; Pred. No. 0.00088;  
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;  
QY 15 EVIEGPNATVLKGSQARFNCTVSQ---WKLIMWALSDMVLSVPMEPIITNDRFTS- 70  
DB 125 EIVDSASELTA--GVNPKVGTCSGYSYAGTLSWELDG-----KPLVNEKGVSV 173  
QY 71 ----QRVDCQGNFT--SEMIHNVPSDSGNIR----CSLQNSRLHGSAYLTVQVMGELF 120  
DB 174 KEQTRRHPTGLFTLQSELM---VTPARGDPRPTSCFSPLPRHRLRTAPIQPRVM 230  
QY 121 IP-----SVNLVVAENEP-----CEVTCPLSHWTRLPDISWELGLLVSHSY 162  
DB 231 EPVPLEEVQLVW---EPFEGAVAPGGTVTLTCEVPAQS-----PQIHWKQ----- 274  
QY 163 YFVPEPSDLQSAVSIILATPQSNGLTLCVATWKS LKARKSATVNLTVIRCPQDTGGINI 222  
DB 275 -GVPLPLPSPVLIILPEIGFQDQGTYS CVATHSHSGFQESRAVISIIE-PGEEG----- 327  
QY 223 PGVLSLPSLGLPTWGVGLGLAGTLLIT-----PTCTLTIRCCRRRCGCCNCC 275

Search completed: November 13, 2003, 03:31:08  
Job time : 36.7643 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 02:52:59 ; Search time 33.7643 Seconds  
(without alignments)  
2919.536 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059

Sequence: 1 MGLVPLHGSGSNEVIEGP.....HPQAFNLASPEKVSNTTVV 382

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1694	82.3	315	4 Q9NS15	Q9ns15 homo sapien
2	575.5	28.0	270	11 Q9D8G2	Q9d8g2 mus musculus
3	152	7.4	492	11 Q8C0U8	Q8c0u8 mus musculus
4	141	6.8	168	11 Q9D912	Q9d912 mus musculus
5	138	6.7	168	11 Q8CH20	Q8ch20 mus musculus
6	136.5	6.6	173	11 Q9DAK2	Q9dak2 mus musculus
7	135	6.6	1496	4 Q92626	Q92626 mus sapien
8	133	6.5	1427	13 Q91562	Q91562 xenopus lae
9	132.5	6.4	344	13 Q9DF61	Q9df61 gallus gall
10	131	6.4	697	11 Q9D290	Q9d290 mus musculus
11	131	6.4	697	11 Q9DBP0	Q9dbp0 mus musculus
12	129.5	6.3	164	5 Q22048	Q22048 caenorhabdi
13	127.5	6.2	315	13 Q9DC15	Q9dci15 gallus gall
14	127.5	6.2	344	13 Q93242	Q93242 gallus gall
15	127.5	6.2	871	11 Q61987	Q61987 mus musculus
16	127.5	6.2	881	11 Q61988	Q61988 mus musculus

17	127	6.2	197	5 Q17641	Q17641 caenorhabdi
18	127	6.2	1445	11 Q63155	Q63155 rattus norv
19	126.5	6.1	869	4 Q15146	Q15146 homo sapien
20	126	6.1	313	13 Q57596	Q57596 gallus gall
21	125.5	6.1	166	5 Q95QY1	Q95qy1 caenorhabdi
22	125.5	6.1	6632	5 Q01761	Q01761 caenorhabdi
23	124	6.0	816	4 Q8NFA5	Q8nfa5 homo sapien
24	124	6.0	1014	4 Q8NFA6	Q8nfa6 homo sapien
25	124	6.0	1030	4 Q8NFA8	Q8nfa8 homo sapien
26	123.5	6.0	536	11 Q8BJE2	Q8bje2 mus musculus
27	123.5	6.0	6620	4 Q96AA2	Q96aa2 homo sapien
28	123.5	5.9	188	5 Q18238	Q18238 caenorhabdi
29	122.5	5.9	455	4 Q9UIR0	Q9uir0 homo sapien
30	122	5.9	344	11 Q8BG33	Q8bg33 mus musculus
31	121.5	5.9	388	4 Q8NC34	Q8nc34 homo sapien
32	121.5	5.9	534	4 Q8NB18	Q8nb18 homo sapien
33	121.5	5.9	1409	13 Q8J127	Q8j127 brachydanio
34	121.5	5.9	1428	13 Q8AY67	Q8ay67 brachydanio
35	120	5.8	551	5 Q8MSN7	Q8men7 drosophila
36	120	5.8	956	5 Q9W4T9	Q9w4t9 drosophila
37	120	5.8	959	5 Q9N9Y9	Q9n9y9 drosophila
38	120	5.8	975	5 Q97174	Q97174 drosophila
39	119.5	5.8	940	4 Q8NFA7	Q8nfa7 homo sapien
40	119	5.8	695	11 Q9JJ07	Q9jj07 rattus norv
41	119	5.8	1340	4 Q8NDA2	Q8nda2 homo sapien
42	118	5.7	275	11 Q9JK39	Q9jk39 mus musculus
43	118	5.7	764	5 Q9W4U1	Q9w4u1 drosophila
44	118	5.7	764	5 Q8MQQ1	Q8mqq1 drosophila
45	117.5	5.7	390	4 Q9H1X9	Q9h1x9 homo sapien

#### ALIGNMENTS

RESULT 1

Q9NS15 Q9NS15 PRELIMINARY; PRT; 315 AA.

AC Q9NS15, 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE IGSF5 protein (Fragment).

GN IGSF5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,

Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,

Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,

Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,

Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,

Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,

Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,

Minoshima S., Shimizu N., Nordstok G., Hornischer K., Brandt P.,

Scharer M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,

Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,

Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,

Lehrach H., Reinhardt R., Yaspo M.L.,

Submitted (WAY-2000) to the EMBL/GenBank/DBJ databases.

EMBL; AL163280; CAB90447.1; -.

DR InterPro; IPR006209; EGF\_like.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003006; IG\_MHC.

DR Pfam; PF00047; ig; 2.

DR SMART; SM00409; IG; 2.

DR PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PSS0835; IG\_LIKE; 2.

FT NON\_TER 1

FT NON\_TER 315

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 12, 2003, 16:40:51 ; Search time 4542.26 Seconds  
(without alignments)  
10582.571 Million cell updates/sec

Title: US-09-729-264-1  
Perfect score: 1175  
Sequence: 1 ctgtctgcccattcgaataa.....gtaatacaactgtagttag 1175

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBank:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rtd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1175	100.0	1175	6	AX380396	AX380396 Sequence
2	1125.4	95.8	2051	9	AK092516	AK092516 Homo sapi
3	1123.4	95.6	1240	6	AX380400	AX380400 Sequence
4	1122.4	95.5	1168	6	AX380398	AX380398 Sequence
5	1067.4	90.8	1139	6	AX380402	AX380402 Sequence
6	381	32.4	1195	6	AX380404	AX380404 Sequence
7	378.8	32.2	1155	10	AF537215	AF537215 Mus muscu
8	320.2	27.3	142742	9	AF121782	AF121782 Homo sapi
9	320.2	27.3	340000	9	HS21C080	AL163280 Homo sapi
10	188.6	16.1	149964	2	AC120145	AC120145 Mus muscu
11	188.6	16.1	181510	2	AC120346	AC120346 Mus muscu
12	188.6	16.1	196900	2	AC020851	AC020851 Mus muscu
13	186	15.8	754	6	AX380408	AX380408 Sequence
14	186	15.8	895	6	AX380406	AX380406 Sequence
15	153.6	13.1	1556	10	BC004806	BC004806 Mus muscu
16	153.6	13.1	47467	2	AC135456	AC135456 Rattus no
17	149.8	12.7	181510	2	AC142238	AC142238 Rattus no
18	149.8	12.7	181510	2	AC120346	AC120346 Mus muscu
19	100	8.5	199665	9	AF064857	AF064857 Homo sapi
20	58	4.9	149964	2	AC120145	AC120145 Mus muscu
21	58	4.9	196900	2	AC020851	AC020851 Mus muscu
22	57	4.9	40205	9	AF045450	AF045450 Homo sapi
23	57	4.9	170121	9	AF064860	AF064860 Homo sapi
24	56.6	4.8	427	9	HS063332	U63332 Human super
25	55.6	4.7	250956	2	AC118358	AC118358 Rattus no
26	55	4.7	1061	9	HSCAGCTG	AJ000501 Homo sapi
27	53	4.5	304785	2	AC118366	AC118366 Rattus no
28	52.8	4.5	110000	3	AE003536_2	Continuation (3 of
29	52.8	4.5	138685	2	AC017582	AC017582 Drosophill
30	52.8	4.5	168249	3	AC093502	AC093502 Drosophill
31	52.8	4.5	185404	3	AC104703	AC104703 Drosophill
32	52	4.4	280780	2	AC125940	AC125940 Rattus no
33	51.8	4.4	397	6	BD010479	BD010479 Diagnosti
34	51	4.3	136171	2	AC141946	AC141946 Rattus no
35	51	4.3	192820	2	AC102163	AC102163 Mus muscu
36	50.6	4.3	2000	6	AX655393	AX655393 Sequence
37	50.6	4.3	147185	9	AC020589	AC020589 Homo sapi
38	50.6	4.3	192292	2	AC129068	AC129068 Rattus no
39	50.6	4.3	320082	2	AC113925	AC113925 Rattus no
40	50	4.3	110000	2	AC108346_1	Continuation (2 of
41	50	4.3	226725	2	AC113770	AC113770 Rattus no
42	50	4.3	237179	2	AC128466	AC128466 Rattus no
43	50	4.3	244322	2	AC123135	AC123135 Rattus no
44	49.8	4.2	299	9	HSAl31896	AJ131896 Homo sapi
45	49.8	4.2	209706	2	AC119611	AC119611 Rattus no

ALIGNMENTS

RESULT 1  
AX380396  
LOCUS AX380396 1175 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 1 from Patent WO0200710.  
ACCESSION AX380396  
VERSION AX380396.1 GI:19575326  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS Welcher, A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.  
TITLE B7-like molecules and uses thereof  
JOURNAL Patent: WO 0200710-A 1 03-JAN-2002;

Angen, Inc. (US)  
Location/Qualifiers  
1. .1175  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
27. .1175  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD28216.1"  
/db\_xref="GI:19575327"  
/translation="MGLVFLHSGSGSNEVIEPQATVILKGSQARFNCVTSQGWKL  
MWALSDMVLSVRPMEPIIINDRFTSQRYDQGNFTSEMIHNVRSDSNIRCSQIN  
SRHGSAYLTQVMGELFIPSNLVVAENEPCEVTLPSHWIRLPDISWELGLVRS  
SYTFVPEPSLDQAGSILATPQSNGLTCTVATWKSILKARKSATVNLTVIRCPQDTGG  
GINPGLVSLPSLGSFPLPWGKVLGLAGTLLPTCTLTIRCCRRRCCGNCNC  
RCFCRRKRGRIFOFKSEKTEKETETESNGNSGVNSDEQKTTDTFASLPKSC  
ESSDPORNSCCPPHORAPRPRASHPQASFNLASPEKVSNTIV"  
BASE COUNT 295 a 307 c 286 g 287 t  
ORIGIN  
Query Match 100.0%; Score 1175; DB 6; Length 1175;  
Best Local Similarity 100.0%; Pred. No. 2e-269;  
Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGTCGCCCATCGAATAACAAGAGATGGGGCTTGTAATTTCTCCACGGTTCTGGGT 60  
Db 1 CTGTCGCCCATCTGAATAACAAGAGATGGGGCTTGTAATTTCTCCACGGTTCTGGGT 60  
QY 61 CTGTAATGAAGTCATAGAGAGGCCCCAGAAATGCAACAGTCCTCAAGAGGCTCCACAGGCTC 120  
Db 61 CTGTAATGAAGTCATAGAGAGGCCCCAGAAATGCAACAGTCCTCAAGAGGCTCCACAGGCTC 120  
QY 121 GCTTCAATGCACCGTCTCCAGAGGCTGGAAAGCTCATCATGTGGGCTCTCAGTGACATGG 180  
Db 121 GCTTCAATGCACCGTCTCCAGAGGCTGGAAAGCTCATCATGTGGGCTCTCAGTGACATGG 180  
QY 181 TGGTGCTAAGCGTCAGGCCCATGAGCCCATCATCACCATGACCGCTTACCTTACCTTACA 240  
Db 181 TGGTGCTAAGCGTCAGGCCCATGAGGCCCATCATCACCATGACCGCTTACCTTACCTTACA 240  
QY 241 GGTACGACACAGGGGGAACCTTCACTCGGAGATGATCATCCAAATGTGGAGCCCAAGT 300  
Db 241 GGTACGACACAGGGGGAACCTTCACTCGGAGATGATCATCCAAATGTGGAGCCCAAGT 300  
QY 301 ATTCGGGGAACATCAGATGACGCTCCAGAACAGTCGCGCTGCATGATCTGCTTACCTTA 360  
Db 301 ATTCGGGGAACATCAGATGACGCTCCAGAACAGTCGCGCTGCATGATCTGCTTACCTTA 360  
QY 361 CCGTCCCAAGTTATGGGAGAGCTGTTCAATCCAGTGTAACTTTGTAGTCGCTGAGATG 420  
Db 361 CCGTCCCAAGTTATGGGAGAGCTGTTCAATCCAGTGTAACTTTGTAGTCGCTGAGATG 420  
QY 421 AACCTTGTGAAGTTACTTGTCTACCTTCACACTGACCCGCGCTCCCGGATATTTCTGGG 480  
Db 421 AACCTTGTGAAGTTACTTGTCTACCTTCACACTGACCCGCGCTCCCGGATATTTCTGGG 480  
QY 481 AGCTCGGTCTCCGTGACGCAATTAAGCTATTATTTGTCGGAGCCCGACGACCTTC 540  
Db 481 AGCTCGGTCTCCGTGACGCAATTAAGCTATTATTTGTCGGAGCCCGACGACCTTC 540  
QY 541 AAAGTGAGTCAGCATCTGCTGCTGACCCACAGAGCAATGGGACTTTGACTTGGCTGG 600  
Db 541 AAAGTGAGTCAGCATCTGCTGCTGACCCACAGAGCAATGGGACTTTGACTTGGCTGG 600  
QY 601 CTACTGTGAAGAGCCTGAAGCCCGCAAGTCTGCAACTGTGAATCTCACTGTGATTCGGT 660  
Db 601 CTACTGTGAAGAGCCTGAAGCCCGCAAGTCTGCAACTGTGAATCTCACTGTGATTCGGT 660  
QY 661 GTCCCCAAGACACTGGAGGTGGTATTATATTCAGGTGTTATATCAAGTTTACCGAGTT 720  
Db 661 GTCCCCAAGACACTGGAGGTGGTATTATATTCAGGTGTTATATCAAGTTTACCGAGTT 720

QY 721 TAGGTTTTCAATTGCCTACTTGGGGCAAAAGTTGGACTTGGACTAGCAGGCAACCATGCTTC 780  
Db 721 TAGGTTTTCAATTGCCTACTTGGGGCAAAAGTTGGACTTGGACTAGCAGGCAACCATGCTTC 780  
QY 781 TGACGGCGAGCTGACTCTTACAATACGCTGCTGCTGCTCCGCGCTGCTGCTGCTGCTGCT 840  
Db 781 TGACGGCGAGCTGACTCTTACAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 GCAACTGCTGCTGCCGTTGTTCTTCTGCTGTAGAGAAAAAGAGGATTTCTGATTTCAAT 900  
Db 841 GCAACTGCTGCTGCCGTTGTTCTTCTGCTGTAGAGAAAAAGAGGATTTCTGATTTCAAT 900  
QY 901 TTCAAAAGAAATCTGAAAAAGAGAGACAAACAAAGAAATCTGACAGAAAGTGGAAATG 960  
Db 901 TTCAAAAGAAATCTGAAAAAGAGAGACAAACAAAGAAATCTGACAGAAAGTGGAAATG 960  
QY 961 AAAAATCCGCGCTCAATTCAGATGAACAAAGAACCCAGACACCGCTTCTCTCCCTCCCA 1020  
Db 961 AAAAATCCGCGCTCAATTCAGATGAACAAAGAACCCAGACACCGCTTCTCTCCCTCCCA 1020  
QY 1021 AATCCTGTGAATCCAGTGATCCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCCACGAG 1080  
Db 1021 AATCCTGTGAATCCAGTGATCCTGAAACAAAGAAACAGTAGCTGTGGCCCTCTCCACGAG 1080  
QY 1081 GGGCTGATCAAGCTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCTTTAATCTGGCCA 1140  
Db 1081 GGGCTGATCAAGCTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCTTTAATCTGGCCA 1140  
QY 1141 GTCTGAGAAGGTTCAGTAATACAACTGTAGTATAG 1175  
Db 1141 GTCTGAGAAGGTTCAGTAATACAACTGTAGTATAG 1175  
RESULT 2  
AK092516 2051 bp mRNA linear PRI 15-JUL-2002  
LOCUS Homo sapiens cDNA FLJ35197 fis, clone PLACE6017788, highly similar  
DEFINITION to IGSF5.  
ACCESSION AK092516  
VERSION AK092516.1 GI:21751130  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,  
Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.  
and Isogai, T.  
NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2051)  
AUTHORS Isogai, T. and Yamamoto, J.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
FEATURES  
source  
1. .2051  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db_xref="taxon:9606"		1101		AAAGTG3AAATGAAAACCTCCGGCTACAATTTCAGATGAACAAAGAACCACAGAAAACGGCTT		1160	
/clone="PLACE6017788"		1009		CTCTCCCTCCCAAAATCTCTGTGAATCCAGTGTATCTCTGAACAAAGAAACAGTAGCTGTGGCC		1068	
/tissue_type="placenta"		1161		CTCTCCCTCCCAAAATCTCTGTGAATCCAGTGTATCTCTGAACAAAGAAACAGTAGCTGTGGCC		1220	
/clone_lib="PLACE6"		1069		CTCTCCACAGCGGGCTGATCAACGTCCACCCAGCGCCAGCAAGTCAATCCAGGCTTCTT		1128	
/note="Cloning vector: pME18SFL3"		1221		CTCTCCACAGCGGGCTGATCAACGTCCACCCAGCGCCAGCAAGTCAATCCAGGCTTCTT		1280	
BASE COUNT	580 a 463 c 474 g 534 t	1129		TTAATCTGGCCAGTCTGAGAAGCTCAGTAATACTAGTAGTAG		1175	
ORIGIN		1281		TTAATCTGGCCAGTCTGAGAAGCTCAGTAATACTAGTAGTAG		1327	
RESULT 3							
AX380400		AX380400		1240 bp		DNA	
LOCUS		Sequence 5 from Patent WO0200710.				linear	
DEFINITION		AX380400				PAT 18-MAR-2002	
ACCESSION		AX380400					
VERSION		AX380400.1		GI:19575330			
KEYWORDS		Homo sapiens (human)					
SOURCE		Homo sapiens					
ORGANISM		Homo sapiens					
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE		Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.					
JOURNAL		B7-like molecules and uses thereof					
FEATURES		Patent: WO 0200710-A 5 03-JAN-2002;					
source		Amgen, Inc. (US)					
CDS		Location/Qualifiers					
		1..1240					
		/organism="Homo sapiens"					
		/mol_type="genomic DNA"					
		/db_xref="taxon:9606"					
		80..1240					
		/note="unnamed protein product"					
		/codon_start=1					
		/protein_id="CAD28218.1"					
		/db_xref="GI:19575331"					
		/translation="MERHLTVPEAVGSGSGNEVIEGPNATVLKGSQARFNCTVSQ					
		WKLMALSDMVLVSVRPMEPIITNDRETSQYDQGNFTSEMIHNVEPSDSNIRC					
		SLQNSRLHGSAYLTQVMGELIFSVNLVAENEPEVTCPLPSHWTRLPDLSWEIGLL					
		VSHSYFVPEPSDLQSAVSLALTPOSNGLTTCVATWKSARKSATVNLTVIRCPQ					
		DTGGINIPGVLSLPSLGFSLGWGLAGTMLTPTLTIRCCRRRCGCC					
		NCCRCFCFRRKRGFRIOFKSEKENTKETETESGNSGNSDQKTTETASLP					
		PKCESDPEQRNSCGPPHQRADQRPAPRPPASHPQASPNLASPEKVSNTTVV"					
BASE COUNT	318 a 319 c 305 g 298 t						
ORIGIN							
Query Match		95.6%;		Score 1123.4;		DB 6;	
Best Local Similarity		99.9%;		Pred. No. 4.1e-257;		Length 1240;	
Matches 1124;		Conservative 0;		Mismatches 1;		Indels 0;	
Gaps 0;							
QY	51	GTTTCTGGGTCTGGTAAATGAAGTCTATAGAAGCCCCCAGAAATGCAACAGTCTCTGAAGGCG		110			
Db	116	GTTTCTGGGTCTGGTAAATGAAGTCTATAGAAGCCCCCAGAAATGCAACAGTCTCTGAAGGCG		175			
QY	111	TCCAGGCTCGCTTCAACTGCACCGTCTCCAGGCTCGGAGCTCATCATGTGGGCTCTC		170			
Db	176	TCCAGGCTCGCTTCAACTGCACCGTCTCCAGGCTCGGAGCTCATCATGTGGGCTCTC		235			
QY	171	AGTGACATGGTGGTCTAAGCGTCAAGCGTCAAGCGCTAGGCGCCATGAGCGCCATCATCAACATGACCGCTTC		230			
Db	236	AGTGACATGGTGGTCTAAGCGTCAAGCGTCAAGCGCTAGGCGCCATGAGCGCCATCATCAACATGACCGCTTC		295			
QY	231	ACCTCTCAGAGGTACGACCAAGCGGGAACCTTCACTCGAGATGATCATCCACAAATGTG		290			
Db	296	ACCTCTCAGAGGTACGACCAAGCGGGAACCTTCACTCGAGATGATCATCCACAAATGTG		355			
QY	291	GAGCCAGTGTATCGGGGAAACATCAGATGACGCTCCAGAACAGTCCGCTGCATGGATCT		350			

Db	356																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

TITLE	B7-like molecules and uses thereof				
JOURNAL	Patent: WO 0200710-A 3 03-JAN-2002;				
	Angen, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..1168				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
CDS	8..1168				
	/notes="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD28217.1"				
	/db_xref="GI:19575329"				
	/translation="MWAGAMENRDPGSGSGNEVIEGPONARVLKSGQARFNCTVSG				
	WKLIIMWLSDMVLSVRPEMFIITNDRFTSDQGNFTSEMIHNVEFSDSGNIRLC				
	SLQNSLHGSAYLTVQMGELFIPSNLVVAENEPCEVCLPSHWLWLPDIHWELGL				
	VHSSSYFVPEPSDLQSAVILALTPQSNGLTCVATKSLKARKATVNLTVIRCPQ				
	DTGGGNIIPGLSSLPISLGLPTWKGVLGLAGTLLPTCLTTRCCCRRCRCG				
	NCCRCFCFCRRKRGFRIQFKSEKENTKETEESGNENSCVNSDEQKTTDTASLP				
	PKSCSSDPEQRNSSCGPHQRADQRPAPASHPQASFNLASPEKVSNTTVV"				
BASE COUNT	299 a	302 c	288 g	279 t	
ORIGIN					
Query Match	95.5%;	Score 1122.4;	DB 6;	Length 1168;	
Best Local Similarity	99.5%;	Pred. No. 7e-257;			
Matches 1126;	Conservative	0;	Mismatches	6;	Indels 0; Gaps 0;
Qy	44	CTTCCACGGTTC	TGGGTCTGGTAATGAAGTCATAGAAGGCCCCAGAAATCAACAGTCCT	103	
Db	37	CCCACCGGTTCT	TGGGTCTGGTAATGAAGTCATAGAAGGCCCCCAAAATCAAGAGTCCT	96	
Qy	104	GAAGGGCTCC	AGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTG	163	
Db	97	GAAGGGCTCC	AGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTG	156	
Qy	164	GGCTCTCAGTGA	CATGTGTGTGTGTAGCGTCAGGCCCATGAGGCCCATCATCAATGA	223	
Db	157	GGCTCTCAGTGA	CATGTGTGTGTGTAGCGTCAGGCCCATGAGGCCCATCATCAATGA	216	
Qy	224	CCGCTTCACTCT	CAGAGGTACGACAGGCGGGAACTTCACCTCGGAGATGATCATCCA	283	
Db	217	CCGCTTCACTCT	CAGAGGTACGACAGGCGGGAACTTCACCTCGGAGATGATCATCCA	276	
Qy	284	CAATGTGGAGCC	CAAGTATTCGGGGAACATCAGATGAGCCTCCAGAACAGTCGCTGCA	343	
Db	277	CAATGTGGAGCC	CAGTATTCGGGGAACATCAGATGAGCCTCCAGAACAGTCGCTGCA	336	
Qy	344	TGGATCTGCTT	ACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAACT	403	
Db	337	TGGATCTGCTT	ACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAACT	396	
Qy	404	TGTAGTCGTG	AGAAAGAACTTGTGAAGTTACTTGTCTACCTCACTGGACCCGGCT	463	
Db	397	TGTAGTCGTG	AGAAAGAACTTGTGAAGTTACTTGTCTACCTCACTGGACCCGGCT	456	
Qy	464	CCCGGATATTT	CTGGAGCTCGGTCTCCGTCAGCCATTCAAGCTATTATTTTGTTC	523	
Db	457	CCCGGATATTT	CTGGAGCTCGGTCTCCGTCAGCCATTCAAGCTATTATTTTGTTC	516	
Qy	524	GGAGCCCAAGC	CACTTCAAAGTCAGTGAGCATCTCTGGCTCTGACCCCAAGAGCAATGG	583	
Db	517	GGAGCCCAAGC	CACTTCAAAGTCAGTGAGCATCTCTGGCTCTGACCCCAAGAGCAATGG	576	
Qy	584	GACTTTGACTT	GGTGGCTACCTGGAAAGAGCTGAAAGCCCGCAAGTCTGCAACTGTAAA	643	
Db	577	GACTTTGACTT	GGTGGCTACCTGGAAAGAGCTGAAAGCCCGCAAGTCTGCAACTGTAAA	636	
Qy	644	TCTCACTGTG	ATTCGGGTCCCAAGACACTGGAGGTGGTATTAATATTCAGGTGATT	703	
Db	637	TCTCACTGTG	ATTCGGGTCCCAAGACACTGGAGGTGGTATTAATATTCAGGTGATT	696	
Qy	704	ATCAAGTTTAC	CGAGTTTAGTGTTCCTTTCCTTCTGCTTCTGGGCAAGTGTGACTTGGACT	763	

Db	697	ATCAAGTTTACCGAGTTTAGGTTTTTCATTGCGCTACTTTGGGGCAAGAGTTGAGCTTGGA	756
Qy	764	AGCAGGACCATGTTCTGACGCGCAGCGTGTACTTCTTACAAATACGCTGCTGCTGCGG	823
Db	757	AGCAGGACCATGTTCTGACGCGCAGCGTGTACTTCTTACAAATACGCTGCTGCTGCGG	816
Qy	824	CCGTCGTTGTTGGCTGCAACTGCTGCTGCGGTTGTTGTTCTGCTGTAGAGAAAAG	883
Db	817	CCGTCGTTGTTGGCTGCAACTGCTGCTGCGGTTGTTGTTCTGCTGTAGAGAAAAG	876
Qy	884	AGGATTTTCGTTATCAATTTCAAAAGAAATCTGAAAAGAGAGAGACAAACAAAGAACTGA	943
Db	877	AGGATTTTCGTTATCAATTTCAAAAGAAATCTGAAAAGAGAGAGACAAACAAAGAACTGA	936
Qy	944	GACAGAAAGTGGAAATGAAACTCCGGCTACAAATTCAGATGAAACAAAGAACACAGACAC	1003
Db	937	GACAGAAAGTGGAAATGAAACTCCGGCTACAAATTCAGATGAAACAAAGAACACAGACAC	996
Qy	1004	CGCTTCTCCTCCCAATCCTGTAATCCTGTAATCCTGTAATCCTGTAATCCTGTAATCCTG	1063
Db	997	CGCTTCTCCTCCCAATCCTGTAATCCTGTAATCCTGTAATCCTGTAATCCTGTAATCCTG	1056
Qy	1064	TGGCCCTCCTCACCAGCGGGCTGATCAACGTCCTCACCAGCGCAGCAAGTCATCCACAGGC	1123
Db	1057	TGGCCCTCCTCACCAGCGGGCTGATCAACGTCCTCACCAGCGCAGCAAGTCATCCACAGGC	1116
Qy	1124	TTCTTTTAATCTGCCAGTCTCTGAGAGGTCAGTAATACAACTAGTATAG	1175
Db	1117	TTCTTTTAATCTGCCAGTCTCTGAGAGGTCAGTAATACAACTAGTATAG	1168
RESULT 5			
AX380402			
LOCUS			
DEFINITION			
AX380402			
ACCESSION			
VERSION			
AX380402.1			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Welcher, A. A., Sarmiento, U. M., Schultz, H. J. and Chute, H. T.			
B7-like molecules and uses thereof			
Patent: WO 0200710-A 7 03-JAN-2002;			
Angen, Inc. (US)			
FEATURES			
source			
1..1139			
Location/Qualifiers			
1..1139			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
1..1134			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAD28219.1"			
/db_xref="GI:19575333"			
/translation="MVAGAMENRDPGSGSGNEVIEGPQARVLKGSQARFNCTVSQG			
WKLIMLWSDMVLVSRMPEFIITNDRFTSQRYDQGNLTSEMILHNVEPDSGNIR			
SLQNRJHGSAYLTQVMYELFIPSVNLVVAENPECEVTLPSHWTRLPDISWELGLL			
VSHSYFVPEPSDQSAVSLTALPQSNGLTCTVATWKSILKARKSATVNLTVIRCPQ			
DTGGINIPGVLSLPSLPSLPTWKGVLGLAGTLLTPTCTLTIRCCCRRCRCG			
NCCRCFCRRKRGKLRKROTKLRQKVMKTPATONKRPQTPLLSLPFPNVP			
ILNKETVAVALLTSGLINVHPGQVHRLILLIWFVLRSLQL"			
BASE COUNT			
290 a 300 c 283 g 266 t			
ORIGIN			
Query Match			
Best Local Similarity 90.8%; Score 1067.4; DB 6; Length 1139;			
Matches 1103; Conservative 0; Mismatches 6; Indels 22; Gaps 1;			
Qy	44	CCTCCAGGTTCTGGGTCTGTGTAATGAGTCAATAGAGGCCGCCAGAAATGCAAGTCCT	103

Db	30	CCACCCCGTTCTGGGTCTGGTAAATGAAAGTCATAGAAAGGCCCCCAAAATGCAAGAGTCCT	89
Qy	104	GAAGGGCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTG	163
Db	90	GAAGGGCTCCAGGCTCGCTTCAACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTG	149
Qy	164	GGCTCTCAGTGAATGTTGGTCTTAAGGCTCAGGCCCATGAGCCCATCATCAACAATGA	223
Db	150	GGCTCTCAGTGAATGTTGGTCTTAAGGCTCAGGCCCATGAGCCCATCATCAACAATGA	209
Qy	224	CCGCTTCACTCTCAGAGGTACGACCGGGCGGAACTTACCTCGGAGATGATCATCCA	283
Db	210	CCGCTTCACTCTCAGAGGTACGACCGGGCGGAACTTACCTCGGAGATGATCATCCA	269
Qy	284	CAATGTGGAGCCCAAGTGAATTCGGGGAAACATCAGATGAGCCCTCCAGAACAGTCGCTGCA	343
Db	270	CAATGTGGAGCCCAAGTGAATTCGGGGAAACATCAGATGAGCCCTCCAGAACAGTCGCTGCA	329
Qy	344	TGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCAGTGTAAATCT	403
Db	330	TGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCAATCCAGTGTAAATCT	389
Qy	404	TGTAGTCGCTGAGAAATGAACCTTGTGAAGTTTACTTGTACCTCACAACCTGACCCGGCT	463
Db	390	TGTAGTCGCTGAGAAATGAACCTTGTGAAGTTTACTTGTACCTCACAACCTGACCCGGCT	449
Qy	464	CCCGATATTTCTCGGAGCTCGGTCTCTGGTTCAGCCATTCAAGCTATTAATTTGTTC	523
Db	450	CCCGATATTTCTCGGAGCTCGGTCTCTGGTTCAGCCATTCAAGCTATTAATTTGTTC	509
Qy	524	GGAGCCAGGACCTTCAAAGTGCAGTGCAGATCCTGGCTCTGACCCACAGAGCAATGG	583
Db	510	GGAGCCAGGACCTTCAAAGTGCAGTGCAGATCCTGGCTCTGACCCACAGAGCAATGG	569
Qy	584	GACTTTGACTTGGCTGCTTACCTGGAAGAGCTTCAAGCGCCGCAAGTCTGCAACTGTA	643
Db	570	GACTTTGACTTGGCTGCTTACCTGGAAGAGCTTCAAGCGCCGCAAGTCTGCAACTGTA	629
Qy	644	TCTCACTGTGANTCGGTGTCCTCCCAAGACACTGAGGTGGTATTAATATTCAGGTGTA	703
Db	630	TCTCACTGTGANTCGGTGTCCTCCCAAGACACTGAGGTGGTATTAATATTCAGGTGTA	689
Qy	704	ATCAAGTTTACCGAGTTAGTGTTCATTGTCCTACTTGGGCAAGTTGAGCTTGGACT	763
Db	690	ATCAAGTTTACCGAGTTAGTGTTCATTGTCCTACTTGGGCAAGTTGAGCTTGGACT	749
Qy	764	AGCAGGCACCATGCTTCTGACGCGCAGCTGCTTACAAATACGCTGCTGCTGCTGCGG	823
Db	750	AGCAGGCACCATGCTTCTGACGCGCAGCTGCTTACAAATACGCTGCTGCTGCTGCGG	809
Qy	824	CCGTCGTTGTTGGCTGCAACTGCTGCTGCGGTGTTGTTTCTGTGTAGAGAAAAG	883
Db	810	CCGTCGTTGTTGGCTGCAACTGCTGCTGCGGTGTTGTTTCTGTGTAGAGAAAAG	869
Qy	884	AGGATTTTCGTTATCAATTTCAAAGAAATCTGAAAAGAGAGAGACAAACAAAGAACTGA	943
Db	870	AG-----GAAATCTGAAAAAGAGAGAGACAAACAAAGAACTGA	907
Qy	944	GACAGAAAGTGGAAATGAAACTCCGGCTCAATTCAGATGAAACAAAGACACAGACAC	1003
Db	908	GACAGAAAGTGGAAATGAAACTCCGGCTCAATTCAGATGAAACAAAGACACAGACAC	967
Qy	1004	CGCTTCTCCTCCCAATTCCTGTGAATCAAGTGTCTGAAACAAAGAAACAGTAGCTG	1063
Db	968	CGCTTCTCCTCCCAATTCCTGTGAATCAAGTGTCTGAAACAAAGAAACAGTAGCTG	1027
Qy	1064	TGGCCCTCCTCAGCAGCGGCTGATCAACGTCACCCAGGCCAGAGTCATCCACAGGC	1123
Db	1028	TGGCCCTCCTCAGCAGCGGCTGATCAACGTCACCCAGGCCAGAGTCATCCACAGGC	1087
Qy	1124	TTCTTTTAATCTGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATA	1174
Db	1088	TTCTTTTAATCTGCCAGTCTCTGAGAGGTCAGTAATACAACTGTAGTATA	1138

```
RESULT 6
AX380404
DEFINITION Sequence 9 from Patent WO0200710.
ACCESSION AX380404
VERSION AX380404.1 GI:19575334
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Welcher, A.A., Samiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE B7-like molecules and uses thereof
JOURNAL Patent: WO 0200710-A 9 03-JAN-2002;
Amgen, Inc. (US)
FEATURES
source Location/Qualifiers
1..1195
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
53..1165
/note="unlabeled protein product"
/codon_start=1
/protein_id="CAD28220.1"
/db_xref="GI:19575335"
/translation="MEGSRDVLAVILVLAQLTASGSSVQIIEGPQNVTVLKDSBAHF
NCTVPHGKLLMTWNTQVLSLTTOGPIITNNRTVSYNSTDSFELIHDVQPS
DSGVSQCSLONSHGSGFASLQVQMGTLNPSNNLIVTEGEPNCVTCVAGWTSLPDI
SWELEVPVSHSYNSFLPEGNFMRVLSVLDLTPNGTLTCVAELKDLQASLSLVNL
TVQPPSPSIEGEPALPTWAILLJAVAFSLLLIILVLIIFPCCOASRSREKSTY
QNEIRKSANMRNKADPETKLKSGKENYGYSDKAQAQTASLPKPSAEVSLPERKSS
SLPYOLNKHQGPATHPRVSPDIAPOKVRNVTLV"
BASE COUNT 313 a 304 c 291 g 287 t
ORIGIN
Query Match 32.4%; Score 381; DB 6; Length 1195;
Best Local Similarity 61.9%; Pred. No. 5.2e-80;
Matches 712; Conservative 0; Mismatches 370; Indels 69; Gaps 4;
QY 24 GAGATGGGCTTGTGATTTCTCCACGGTCTGGCTGGTGGTAAAGTCAATAGAGGC 83
DB 83 GTGCTGGTCACTGCTGCTGAGTGAAGCTTCGGATCCAGTATCAGATCATAGAGGT 142
QY 84 CCCCAGATGCAACAGTCTCTGAAGGCTCCACGGCTCGCTTCAACTGCAACCGTCTCCAG 143
DB 143 CCTCAGATGTAACAGTCTAAAGACTCAGAGGCTCACTTCAACTGCAACCGTCTCAC 202
QY 144 GGTGGAAGCTCATCATGTGGGCTCTCAGTACATGGTGTGCTAAGCGTCAAGGCCCATG 203
DB 203 GGTGGAAGCTTCTCATGTGGTCTTAAACCAATGGTGTGCTGAGTCTCACACCCAA 262
QY 204 GAGCCATCATCAACCAATCAGCTTCACTCTCAGAGGTACACAGCGGGGGAATTC 263
DB 263 GGACCATCATCAACCAACCGCTTCACTATGCGAGTACACAGCACTGACAGCTTC 322
QY 264 ACCTCGAGATGATCATCCCAATGTGGAGCCAGTGATTTCGGGGGAACATCAGATCAGC 323
DB 323 ATCTCGAGTGTGATCATCCATGATGTCAGCCAGTGAAGTCTCGGATCCCGTGAATGAGC 382
QY 324 CTCAGAACAGTCCGCTGATGATGCTGTACCTTACGTCACAGTATGGAGAGCTG 383
DB 383 CTCAGAACAGCATGGTGTGATGCTGCTTCCCTCTCAGTGCAGATGATGGGACCCCTG 442
QY 384 TTCAATCCAGTGTAACTTGTAGTCGCTGAGAATGAACCTTGTGAAGTACTTGTCTA 443
DB 443 AACATTCCTAGCAACACCTTATAGTCACTGAGGGTGAACCTTGTATGTGACTTGTCTAT 502
QY 444 CCTCACTAGGACCGGCTCCCGGATATTTCTCGGAGCTCGGTCTCTCTGCTGAGCCAT 503
DB 503 GCGTGGGCTGGACCTCACTCCCGGATATTTCTCGGAGCTTGAAGTTCCTGTAAGCCAT 562
```

```
QY 504 TCAAGCTATTATTTGTTCCGAGCCCGAGCACTTCAAAGTGCAGTGCAGTATCTGGT 563
DB 563 TCGAGTTACAATCTCTTTCTGAGCCGGCAACTTATGAGGGTCTTGAAGTCTCTGGAC 622
QY 564 CTGACCCACAGAGCAATGGGACTTTGACTTCGCTGGCTTACCTGGAAGAGCTGAGGCC 623
DB 623 CTCACCACTGGGCAACGGGACCTTGACTTGTGTGGGAGAGCTGAGAGACTTGCAGGCC 682
QY 624 CGCAAGTCTGCAACTGTAAATCTCACTGTGATTCCCGTGTCCCAAGACACTCGAGTGT 683
DB 683 AGCAAGTCTTAACTGTCACTGACTGTGTTGAGCTTCCACCTGACAGTATTGGAGAG 742
QY 684 ATTAATATTCAGGTGTATTATCAAGTTTACCGAGTTTAGGTTTTTCAATGCTCTACTGG 743
DB 743 GA-----AGGCCAGCACTGCGGACCTGG 766
QY 744 GCAAGTTGGACTTGGACTAGCAGGACCACTGCTTCTGACCGCCGAGTGTACTCTTACA 803
DB 767 GCCATCATCTGCTGCGGAGTGGCCCTTTCTTGTGCTTCTGATCCTGATCAATTTTGTAT 826
QY 804 ATACGCTGCTGCTGCTGCGCGCTGCTGTTGTTGTGCTGCAACTGCTGCTGCGGTTGTGT 863
DB 827 ATAATATTTCTGTTGCTGTTGT-----GCTCCAGGAGAGAAAGGAATC 873
QY 864 TTCCTGCTGAGAGAAAAAGAGGATTTCGTATTCATTTCAAAGAAATCTGAAAAAGAG 923
DB 874 TACTTATCAAAATGAATAAGGAAATCTGCAACATGAGGACAA-----AC 919
QY 924 AAGACAAACAAAGAACTGAGACAGAAATGGAATGAAATGAACTCCGGCTACAAATTCAGAT 983
DB 920 AAGCAGATCCCGAGACAAAGTTAAAGAGTGAAGAAAGTACGGGTACAGTTCGAT 979
QY 984 GAACAAAAGACACAGACACCGCTTCTCTCCCTCCCAATCTCTGTAATCCAGTGTATCT 1043
DB 980 GAGCAAGGCTGCACAGACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1039
QY 1044 GAACAAAAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
DB 1040 GAAAAACGAGCAGTA-----GCTTCTCTTATCAGGAACTCAATAAATCAGCCCGGT 1093
QY 1104 CCAGCAAGTCACTCCACAGGCTTCTTTAATCTGGCCAGTCTCTGAGAGTCAAGTAAACA 1163
DB 1094 CCAGCAACTCACTCCACAGGCTTCTTTGATCTGCACTGCACTGCACTGCACTGCACTG 1153
QY 1164 ACTGTAGTATA 1174
DB 1154 ACTTTAGTGA 1164
RESULT 7
AF537215 1155 bp mRNA linear ROD 03-JUN-2003
LOCUS Mus musculus cell adhesion molecule JCAM mRNA, complete cds.
DEFINITION AF537215
ACCESSION AF537215
VERSION AF537215.1 GI:31339733
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1155)
REFERENCE Hirabayashi, S., Tajima, M., Yao, I., Nishimura, W., Mori, H. and
AUTHORS Hata, Y.
TITLE JAM3, a Junctional Cell Adhesion Molecule Interacting with a Tight
Junction Protein, MAGI-1
JOURNAL Mol. Cell. Biol. 23 (12), 4267-4282 (2003)
MEDLINE 22658521
PUBMED 12773569
REFERENCE 2 (bases 1 to 1155)
AUTHORS Hata, Y., Hirabayashi, S. and Tajima, M.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2002) Medical Biochemistry, Tokyo Medical and
```

Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan  
Location/Qualifiers  
1. .1155

/organism="Mus musculus"  
/mol\_type="mrna"  
/strain="Swiss Webster"  
/db\_xref="taxon:10090"

43. .1155

/note="interacts with MAGI-1; localized at tight junctions in epithelial cells"

/codon\_start=1

/product="cell adhesion molecule JCAM"

/protein\_id="AAP49218.1"

/translation="MBSGRDLVLAIVLAQTASGSYQIIEGPQNVTVLKDEAHF

NCVTHGKLMWLNQWVLSLTTOGPIITNNEFTVASTNSDTSFSELIHVDQPS  
DSGVQSLQNSGFGAFSLVQVNGTLNIPNNLIVTEGPCNVTCVAVGWTSLPDI  
SWELEVPVSHSYNSFLSEGNFMRVLSVLDLTPLGNGTLCTVAELKDOIQASKSLTVNL  
TVVPPDPSIGEEGPALPTWAILLLAVAFSLLLLIIVLIIIFCCCAASRREKESTY  
QNEIRKANMTNKADPETKUGKRENTYSSDEAKAAQTASLPKSAEVSLEPKRSS  
SLPYOELNKHQPGPAPHRVSRFDIASPKQVENVTLV"

BASE COUNT 299 a 296 c 283 g 277 t

ORIGIN

Query Match 32.2%; Score 378.8; DB 10; Length 1155;

Best Local Similarity 61.7%; Pred.No. 1.7e-79;

Matches 711; Conservative 0; Mismatches 372; Indels 69; Gaps 4;

QY 24 GAGATGGGGCTGTGATTTCTCCTCACCGGTTCTGGGCTGTGTAATGAAGTCATAGAAGGC 83  
DB 73 GTGCTGTATCCTCGGCTCAGTCGACAGCTTCGGATCCAGTTATCAGATCATAGAAGGT 132  
QY 84 CCCAGAAATCAACAGTCTGAAGGCTCCAGGCTCGCTTCAACTGACCGTCTCCAG 143  
DB 133 CCTCAGAATGTAACAGTCTTAAGGACTCAGAGGCTCACTTCACTGACCGTGACTCAC 192  
QY 144 GGCTGGAAGCTCATCATGTGGCTCTCAGTCGACATGGTGGTGAAGCGTCAGGCCCATG 203  
DB 193 GGCTGGAAGCTCTCATGTGAGCTTTAACCAGATGGTGGTGGTGGTGGTGGTGGTGGT 252  
QY 204 GAGCCCATCATCACCAGTACCGCTTCACTCTCAGAGGTACGACGCGGGGGGAACTTC 263  
DB 253 GGACCCATCATCACCAGTACCGCTTCACTCTCAGAGGTACGACGCGGTTCAGAGCTTC 312  
QY 264 ACCTCGAGATGATCATCAAGTGTGAGCCGAGTATTCGGGAAATCAGATGACG 323  
DB 313 ATCTCGAGTGTGATCATCATGATGTGAGCCGAGTATTCGGGAAATCAGATGACG 372  
QY 324 CTCCAGAACAGTCGCTCATGATGCTCTTACCTTACCTTACCTTACCTTACCTTACCTT 383  
DB 373 CTGAGAACAGCCATGGGTTGGATCTGCTTCTCCTCAGTGCAGATCATGGGACCCCTG 432  
QY 384 TTCAATCCAGTGTAACTTGTAGTCTGCTGAGATGAACCTTGTGAAGTTACTTGTCTA 443  
DB 433 AACATTCCTAGCAACACCTTATAGTCACTGAGGTGAACCTTGTGAAGTTACTTGTCTA 492  
QY 444 CCCTCAGACTGGACCGGCTCCGAGATTTCTGGGAGCTCGTCTCTCTGTCAGCCAT 503  
DB 493 GCCGTGGGCTGGACCTCACTCCCGGATATTTCTTGGGAGCTTGAGGTTCCCGTAAGCCAT 552  
QY 504 TCAAGCTATTATTTTGTTCGGAGCCGACGACCTTCAAGTGCAGTGCAGTGCATCTGCT 563  
DB 553 TCGAGTTACATTCCTTCTGAGTCTGGGACACTTATGAGGTTCTGAGTCTCTGAC 612  
QY 564 CTGACCCCAAGAGCAATGGAGCTTTGACTTGGTGGCTACCTGGAAGAGCTGAAGGCC 623  
DB 613 CTCACACCTGGGCAAGGACCTTGACTTGTGGCAGAGCTCAAGGACTTGCAGGCC 672  
QY 624 CGCAAGCTGCAACTGTAAATCTCACTGTGATTCGGTGTCCCAAGACAGCTGGAGGTGGT 683  
DB 673 AGCAAGCTCTTAATCTGTCAACCTGACTGTGTTGAGCTCCACCTGACGATTTGGAGAG 732  
QY 684 ATTAATATCCAGGTGTATTATCAAGTTTACCGAGTTTAGGTTTTTCTTCTGCTACTTGG 743

DB 733 GA-----AGGCCAGCAGCTGCCGACCTGG 756  
QY 744 GGCAAGATTGGACTTGGACTAGCAGGACCATGCTTCTGAGCCGACGTTACTCTTACA 803  
DB 757 GCCATCATCTGCTGGCAGTGGCCTTTCTCTTGTGCTGCTGATCATGTTTGTATT 816  
QY 804 ATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863  
DB 817 ATAATATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863  
QY 864 TTCTGCTGTAG 923  
DB 864 TACTTATCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 909  
QY 924 AAGCAAAACAAAGAACTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983  
DB 910 AAGCAGATCCGGAGACAG 969  
QY 984 GAACAAAGACCAAGACAGACAGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1043  
DB 970 GAGCAAGGCTGCACAGACTGCT 1029  
QY 1044 GAACAAAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103  
DB 1030 GAAACAGCAGCAGTA-----GCCCTTCTTATCAGGAACCTCAATAAATCAGCCCGGT 1083  
QY 1104 CCAGCAAGTATCACAGGCTTCTTTTAAATCTGGCAGTCTCTGAGAGGTCAGTAATACA 1163  
DB 1084 CCAGCAACTATCACGGGTTCTTTTACATGCGCAGTCTCTGAGAGTCTCAGAGGTCAGAAATGTG 1143  
QY 1164 ACTGTAGTATAG 1175  
DB 1144 ACTTATGTAG 1155  
RESULT 8  
AF121782 142742 bp DNA linear PRI 02-FEB-1999  
LOCUS Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.  
DEFINITION AF121782  
ACCESSION AF121782  
VERSION AF121782.1 GI:4210991  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 142742)  
AUTHORS Taudien,S., Dagand,B., Hildmann,T., Nordsiek,G., Drescher,B.,  
Schattevoy,R., Weber,J., Schilling,M., Menzel,U., Yaspo,M.L. and  
Rosenthal,A.  
TITLE Direct Submission  
JOURNAL Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
FEATURES  
source  
1. .142742  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clones="PAC 206A10"  
3. .118  
/evidence=not experimental  
/rpt\_family="WIR"  
repeat\_region complement(140..415)  
/evidence=not experimental  
/rpt\_family="MLT2B"  
repeat\_region complement(668..881)  
/evidence=not experimental  
/rpt\_family="WERS8A"  
exon complement(1078..1162)  
/note="GRAIL, score = 86.000%, comment = excellent shadow"

```

/evidence=not_experimental
1114. .1403
/evidence=not_experimental
/rpt_family="MLT1I"
repeat_region
1481. .1603
/evidence=not_experimental
/rpt_family="L2"
exon
1576. .1710
/note="M2EF, score = 66.1%"
/evidence=not_experimental
1576. .1710
/note="GRAIL, score = 99.000%, comment = excellent"
/evidence=not_experimental
complement(1693. .1741)
repeat_region
/rpt_family="MIR"
repeat_region
complement(1780. .2326)
/evidence=not_experimental
/rpt_family="MLT1E"
exon
complement(1834. .1955)
/note="GRAIL, score = 50.000%, comment = good shadow"
/evidence=not_experimental
2629. .2831
repeat_region
/rpt_family="AluSx"
repeat_region
3012. .3122
/evidence=not_experimental
/rpt_family="MLT1J"
repeat_region
3123. .3478
/evidence=not_experimental
/rpt_family="THE1B"
repeat_region
3479. .3618
/evidence=not_experimental
/rpt_family="MLT1J"
exon
complement(3502. .3540)
/note="GRAIL, score = 65.000%, comment = good shadow"
/evidence=not_experimental
4674. .4781
exon
/note="GRAIL, score = 42.000%, comment = marginal"
/evidence=not_experimental
4712. .4749
exon
/note="Xpound exon prediction, score = 70% (0%)"
/evidence=not_experimental
5385. .6104
repeat_region
/evidence=not_experimental
/rpt_family="HERVL"
exon
6682. .6751
/note="M2EF, score = 95.1%"
/evidence=not_experimental
6746. .6751
exon
/note="GenScan, score = 1.83%, comment = Initial_exon 6 bp
frame: 1 phase: 0"
/evidence=not_experimental
6901. .7339
repeat_region
/evidence=not_experimental
/rpt_family="MLT1C"
repeat_region
7347. .7457
/evidence=not_experimental
/rpt_family="LTR16C"
repeat_region
complement(7561. .7696)
/evidence=not_experimental
/rpt_family="MIR"
repeat_region
complement(7819. .8220)
/evidence=not_experimental
/rpt_family="MLT1J"
exon
8097. .8119
/note="GRAIL, score = 45.000%, comment = marginal"
/evidence=not_experimental
complement(8738. .9232)
repeat_region
/evidence=not_experimental
/rpt_family="MLT1C"
exon
complement(9273. .9420)
/note="GRAIL, score = 71.000%, comment = good"

```

```

/evidence=not_experimental
9290. .9371
/note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence=not_experimental
9966. .10041
repeat_region
/note="homology = 100.00%, score = 38, counts = 2"
/evidence=not_experimental
/rpt_type=tandem
/rpt_unit=tctctgcacactctctgttgagaggtcggtttatcg
10244. .10664
/evidence=not_experimental
/rpt_family="MLT1E"
repeat_region
10665. .11028
/evidence=not_experimental
/rpt_family="MLT1F"
repeat_region
11318. .11417
/evidence=not_experimental
/rpt_family="LMC/D"
repeat_region
11503. .11818
/evidence=not_experimental
complement(11525. .11777)
/rpt_family="AluSg"
exon
/note="M2EF, score = 79.4%"
/evidence=not_experimental
complement(12346. .12827)
repeat_region
/evidence=not_experimental
/rpt_family="MLT1C"
repeat_region
12902. .13180
/evidence=not_experimental
/rpt_family="AluSg"
exon
complement(13057. .13161)
/note="M2EF, score = 76.1%"
/evidence=not_experimental
complement(13226. .13364)
repeat_region
/evidence=not_experimental
/rpt_family="LMC2"
repeat_region
complement(13613. .13908)
/evidence=not_experimental
/rpt_family="AluSx"
repeat_region
complement(14317. .14614)
/evidence=not_experimental
/rpt_family="LTR37A"
exon
complement(14962. .15036)
/note="M2EF, score = 93%"
/evidence=not_experimental
15255. .15299
repeat_region
/evidence=not_experimental
/rpt_family="LIPa6"
exon
complement(15469. .15497)
/note="Xpound exon prediction, score = 62% (0%)"
/evidence=not_experimental
16045. .16768
repeat_region
/evidence=not_experimental
/rpt_family="MER21A"
exon
complement(16555. .16655)
/note="GRAIL, score = 54.000%, comment = good"
/evidence=not_experimental
16823. .16850
repeat_region
/evidence=not_experimental
/rpt_family="MER34"
repeat_region
complement(16851. .17208)
/evidence=not_experimental
/rpt_family="THE1A"
repeat_region
17209. .17632
/evidence=not_experimental
/rpt_family="MER34"
repeat_region
complement(17633. .17969)
/evidence=not_experimental
/rpt_family="AluSg1"
repeat_region
complement(18004. .18334)
/evidence=not_experimental
/rpt_family="LTR17"
repeat_region
complement(18335. .18632)

```

```

/evidence=not experimental
/rpt_family="HERV17"
18352 18420
/notes="GRAIL, score = 51.000%, comment = good shadow"
/evidence=not experimental
complement(18397..18584)
/notes="GRAIL, score = 66.000%, comment = good"
/evidence=not experimental
complement(18643..24647)
/evidence=not experimental
/rpt_family="HERV17"
complement(19228..19361)
/notes="WZEF, score = 92.8%"
/evidence=not experimental
20900..21135
/notes="WZEF, score = 51.4%"
/evidence=not experimental
21150..21275
/notes="Xpound exon prediction, score = 74% (0%)"

Query Match 27.3%; Score 320.2; DB 9; Length 142742;
Best Local Similarity 97.6%; Pred. No. 2.2e-65;
Matches 325; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 TTCTCCAGGCTCTGGCTCGTGAATGAGTCATAGAGGCCGCCAGATGCAACAGTC 101
DB 44265 TACCTTCCAGGTTCTGGGCTCGTGAATGAGTCATAGAGGCCGCCAAATGCAAGATC 44324

QY 102 CTGAAGGGCTCCAGGCTCGCTTCAACTGACCGCTCTCCAGGGCTGGAAGCTCATCATG 161
DB 44325 CTGAAGGGCTCCAGGCTCGCTTCAACTGACCGCTCTCCAGGGCTGGAAGCTCATCATG 44384

QY 162 TGGGCTCTCAGTGACATGTTGTGTGCTAGCGTCAGGCCCATGAGCCCATCATCAACAT 221
DB 44385 TGGGCTCTCAGTGACATGTTGTGTGCTAGCGTCAGGCCCATGAGCCCATCATCAACAT 44444

QY 222 GACCGTCTCACTCTCAGAGTACGACAGGCGGGAACCTTCACTTCGAGATGATCATC 281
DB 44445 GACCGTCTCACTCTCAGAGTACGACAGGCGGGAACCTTCACTTCGAGATGATCATC 44504

QY 282 CACAATGTGAGCCCACTGATTGGGGAAACATCAGATGAGCTCCAGAACAGTCGCTG 341
DB 44505 CACAATGTGAGCCCACTGATTGGGGAAACATCAGATGAGCTCCAGAACAGTCGCTG 44564

QY 342 CATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 374
DB 44565 CATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 44597

RESULT 9
HS21C080 340000 bp DNA linear PRI 24-MAY-2000
LOCUS Homo sapiens chromosome 21 segment HS21C080.
DEFINITION AL163280 AP001735 BA000005
ACCESSION AL163280.2 GI:7717369
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hattori.M., Fujiyama.A., Taylor.T.D., Watanabe.H., Yada.T.,
Park.H.S., Toyoda.A., Ishii.K., Totoki.Y., Choi.D.K., Soeda.E.,
Ohki.M., Takagi.T., Sakaki.Y., Taudien.S., Blechschmidt.K.,
Polley.A., Menzel.U., Delabar.J., Kumpf.K., Lehmann.R.,
Patterson.D., Reichwald.K., Rump.A., Schillabel.M., Schudy.A.,
Zimmermann.W., Rosenthal.A., Kudoh.J., Shibuya.K., Kawasaki.K.,
Asakawa.S., Shintani.A., Sasaki.T., Nagamine.K., Mitsuyma.S.,
Antonarakis.S.E., Minoshima.S., Shimizu.N., Nordisk.G.,
Hornischer.K., Brand.P., Schaefer.M., Schoen.O., Desario.A.,
Reichelt.J., Kauer.G., Bloeker.H., Ramser.J., Beck.A., Klages.S.,
Hennig.S., Rieselmann.L., Dagand.E., Wehrmeyer.S., Borzym.K., and
Gardiner.K., Nizetic.D., Francis.F., Lehnach.H., Reinhardt.K., and

```

TITLE  
JOURNAL

## COMMENT

## FEATURES

source

source

source

source

source

Yaspo, M.L.  
Direct Submission  
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing Consortium: \* RIKEN Genomic Sciences Center, Human Genome Research Group \* Institute of Molecular Biotechnology, Genome Analysis \* Keio University School of Medicine, Dept. of Molecular Biology \* GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular Genetics (addresses see below)  
The Chromosome 21 Mapping and Sequencing Consortium consists of:  
\* RIKEN Genomic Sciences Center, Human Genome Research Group, \* Sagami-hara 228-8555, Japan,  
\* e-mail: sakaki@gscc.riken.go.jp  
\* URL: http://hgp.gsc.riken.go.jp/  
and  
\* Institute of Molecular Biotechnology, Genome Analysis, \* Beutenbergstrasse 11, D-07745 Jena, Germany,  
\* e-mail: gscj-submit@genome.imb-jena.de  
\* URL: http://genome.imb-jena.de/  
and  
\* Keio University School of Medicine, Dept. of Molecular Biology, \* Tokyo 160-8582, Japan,  
\* e-mail: shimizu@mb-med.keio.ac.jp  
\* URL: http://adenline.dmb.med.keio.ac.jp/  
and  
\* GBF, Dept. of Genome Analysis,  
\* Mascheroder Weg 1, D-38124 Braunschweig, Germany, \* e-mail: info.genome@gbf.de  
\* URL: http://genome.gbf.de/  
and  
\* Max-Planck Institute for Molecular Genetics,  
\* Ihnestrass 73, D-14195 Berlin, Germany,  
\* e-mail: info-chr21@molgen.mpg.de  
\* URL: http://chr21.rz-berlin.mpg.de/.  
Location/Qualifiers  
1. 340000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
<1..125946  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P160L9, 5' partial"  
/clone\_lib="RPC11, 3-5 PAC library"  
/note="Accession No. AF121897"  
73410..243533  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P70124"  
/clone\_lib="RPC11, 3-5 PAC library"  
/note="Accession No. AF064860"  
224137..2340000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P206A10, 3' partial"  
/clone\_lib="RPC11, 3-5 PAC library"  
/note="Accession No. AF121782"  
286628..2340000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"

```

/clone="BAC-291B3, 3' partial"
/clone_lib="BAC library"
/note="Accession No. AF064857"
1088..1566
/note="LIM4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/note="THE1C"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
join(<3473..3565,12649..12678,15331..15460,18739..19157)
/gene="SH3BGR"
join(3473..3565,12649..12678,15331..15426)
/partial
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CA890445.1"
/db_xref="GI:7717370"
/translators="GSEKAEGETEQAQKEGSDVGNLPAQKNEEGETATEETEE
1AAGAEAEAEAEETAEGBEPEDEDS"
3473..3565
/gene="SH3BGR"
/number=4
repeat_region
5542..5829
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="LIMD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9982)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)

```

```

/note="THE1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12849..12678
/gene="SH3BGR"
/number=5
complement(13115..13425)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(13598..13892)
/note="AluY"

Query Match      27.3%; Score 320.2; DB 9; Length 340000;
Best Local Similarity 97.6%; Pred. No. 2.4e-65;
Matches 325; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 TTCTCCACGGTTCTGGTCTGTAATGAAGTCATAGAGGCCGCCAGAAATGCAACAGTC 101
   |||||
   268445 TACCTTCCAGTTCTGGTCTGTAATGAAGTCATAGAGGCCGCCCAAAATGCAAGAGTC 268504

QY 102 CTGAAGGGTCCCGAGGCTGCTTCACTGACACGCTCCCGAGGCTGGAGAGCTCATCATG 161
   |||||
Db 268505 CTGAAGGGTCCCGAGGCTGCTTCACTGACACGCTCCCGAGGCTGGAGAGCTCATCATG 268564

QY 162 TGGGCTCTCAGTGACATGTTGGTGTAAAGCTCAGGCCCATGAGCCCATCATCAAT 221
   |||||
Db 268565 TGGGCTCTCAGTGACATGTTGGTGTAAAGCTCAGGCCCATGAGGCCCATCATCAAT 268624

QY 222 GACCGCTTCACTCTCAGAGGTACGACACGAGGGGGGAACCTTCACTCGGAGATGATCATC 281
   |||||
Db 268625 GACCGCTTCACTCTCAGAGGTACGACACGAGGGGGGAACCTTCACTCGGAGATGATCATC 268684

QY 282 CACAATGTGAGCCCGAGTATTCGGGGACATCAGATGAGCTCCAGACAGTCCAGACAGTCCGCTG 341
   |||||
Db 268685 CACAATGTGAGCCCGAGTATTCGGGGACATCAGATGAGCTCCAGACAGTCCGCTG 268744

QY 342 CATGATCTGCTTACCTTACCTTACCGTCCCAAGTTATG 374
   |||||
Db 268745 CATGATCTGCTTACCTTACCTTACCGTCCCAAGTTATG 268777

RESULT 10
AC120145
LOCUS
DEFINITION Mus musculus clone RP24-216P24, WORKING DRAFT SEQUENCE, 9 unordered
pieces.
ACCESSION AC120145
VERSION AC120145.3 GI:28893697
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



QY 161 GTGGCTCTCAGTGACATGGTGTGCTGAAGCTCAGGCCCATGAGCCCATCATCAACAA 220  
 |||||  
 Db 89322 GTGGACTCTTAACCAAAATGGTGTGCTGAGTCTCACCAACCAAGGACCATCATCAACAA 89381  
 |||||  
 QY 221 TGACCGCTTACCTCTCAGAGGTACGACCGGGGGGAACCTTCACTCCGCGAGATGATCAT 280  
 |||||  
 Db 89382 CAACCGTTTACCTATGCCAGTTTAAACAGACACTGACAGCTTTCATCTCGGAGTTGATCAT 89441  
 |||||  
 QY 281 CCACAATGTGGAGCCCACTGATTCGGGGAACATCAGATCGAGCTCCAGAACATCGCCCT 340  
 |||||  
 Db 89442 CCATGATGTGAGCCCACTGACTCGGATCGTCAATGCGCTCGCAGAACACCCATGCG 89501  
 |||||  
 QY 341 GCATGGATCTCTTACCTTACCTGCTCCAGTT 371  
 |||||  
 Db 89502 GTTGGATCTGCTTCTCTCAGTGCACAGGT 89532  
 |||||

RESULT 11  
 AC120346  
 LOCUS  
 DEFINITION Mus musculus clone RP23-147E11, WORKING DRAFT SEQUENCE, 15 ordered  
 pieces.  
 AC120346  
 VERSION AC120346.4 GI:24211306  
 KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 181510)  
 Birren,B., Nussbaum,C. and Lander,E.  
 Mus musculus, clone RP23-147E11  
 Unpublished  
 2 (bases 1 to 181510)  
 Birren,B., Linton,D., Nussbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
 Boukhgalter,B., Brown,A., Camarata,J., Campolano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., FitzHugh,M., Gage,D., Galagan,J., Gardyna,S.,  
 Gande,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollata,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 181510)  
 Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,  
 Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,  
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
 Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,

Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,  
 Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (22-OCT-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Oct 22, 2002 this sequence version replaced gi:20514894.  
 All repeats were identified using RepeatMasker:  
 Smit, A.P.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L20921  
 Center clone name: 147 E 11  
 ----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 17801 bases at least Q40  
 Consensus quality: 179340 bases at least Q30  
 Consensus quality: 179851 bases at least Q20  
 Insert size: 176000; agarose-fp  
 Insert size: 180110; sum-of-contigs  
 Quality coverage: 11.3 in Q20 bases; agarose-fp  
 Quality coverage: 11.1 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1 29098: contig of 29098 bp in length  
 \* 29099 29198: gap of 100 bp  
 \* 29199 29864: contig of 666 bp in length  
 \* 29865 29964: gap of 100 bp  
 \* 29965 30244: contig of 280 bp in length  
 \* 30245 30344: gap of 100 bp  
 \* 30345 30971: contig of 627 bp in length  
 \* 30972 31071: gap of 100 bp  
 \* 31072 32648: contig of 1577 bp in length  
 \* 32649 32748: gap of 100 bp  
 \* 32749 37186: contig of 4438 bp in length  
 \* 37187 37286: gap of 100 bp  
 \* 37287 44243: contig of 6957 bp in length  
 \* 44244 44344: gap of 100 bp  
 \* 44344 54253: contig of 9910 bp in length  
 \* 54254 54353: gap of 100 bp  
 \* 54354 70476: contig of 16123 bp in length  
 \* 70477 70576: gap of 100 bp  
 \* 70577 86055: contig of 15479 bp in length  
 \* 86056 86155: gap of 100 bp  
 \* 86156 104095: contig of 17940 bp in length  
 \* 104096 104195: gap of 100 bp  
 \* 104196 123740: contig of 19545 bp in length  
 \* 123741 123840: gap of 100 bp  
 \* 123841 153195: contig of 29355 bp in length  
 \* 153196 153295: gap of 100 bp  
 \* 153296 180206: contig of 26911 bp in length  
 \* 180207 180306: gap of 100 bp  
 \* 180307 181510: contig of 1204 bp in length.

Location/Qualifiers  
 1..181510  
 /organism="Mus musculus"

FEATURES  
 source

```
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-147E11"
/clone_lib="RPCI-23 Female Mouse BAC"
1..29098
/note="assembly_fragment"
clone_end:SP6
vector_side:left
29199..29864
/note="assembly_fragment"
29965..30244
/note="assembly_fragment"
30345..30971
/note="assembly_fragment"
31072..32648
/note="assembly_fragment"
32749..37186
/note="assembly_fragment"
37287..44243
/note="assembly_fragment"
44344..54253
/note="assembly_fragment"
54354..70476
/note="assembly_fragment"
70577..86055
/note="assembly_fragment"
86156..104095
/note="assembly_fragment"
104196..123740
/note="assembly_fragment"
123841..153195
/note="assembly_fragment"
153296..180206
/note="assembly_fragment"
180307..181510
/note="assembly_fragment"
clone_end:T7
vector_side:right
BASE COUNT 49236 a 40873 c 40708 g 49292 t 1401 others
ORIGIN
```

```
Query Match 16.1%; Score 188.6; DB 2; Length 181510;
Best Local Similarity 73.1%; Pred. No. 5.6e-34;
Matches 242; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 41 TTTCTCCAGCTTCTGGGTCGTGTGTAAGTCATAGAGGCCCCAGAAATGCAACAGT 100
Db 152354 TTGCCTTCTAGCTTCCGGATCCAGTTATCAGATCATAGAGGTCCTCAGAAATGTAACAGT 152413

Qy 101 CCTGAAGGCTCCAGGCTCGCTTCAACTGCACGCTCTCCAGGGCTGGAAGCTCATCAT 160
Db 152414 CCTAAGAGCTCAGAGGCTCACTTCAACTGCACGCTGACTCAGGCTGGAGGTTCTCAT 152473

Qy 161 GTGGGCTCTCAGTGACATGGTGTGCTTAAGGCTCAGGCCCATGGAGCCCATCATCAACAA 220
Db 152474 GTGGACTCTTAACCAATGGTGTGCTGAGTCTCACCACCAAGGCCATCATCAACAA 152533

Qy 221 TGRGGCTTCACTCTCAGAGTACGACGAGGGGGAATTCACCTCGGAGTATCAT 280
Db 152534 CAACGGCTTCACTCTCAGAGTACGACGAGGGGGAATTCACCTCGGAGTATCAT 152593

Qy 281 CCACAAATGGAGCCCAAGTATTCGGGGAACATCAGATGCGGCTCCAGAACAGTGCCT 340
Db 152594 CCATGATGTCAGCCCAAGTATTCGGGGAACATCAGATGCGGCTCCAGAACAGTGCCT 152653

Qy 341 GCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 371
Db 152654 GTTTGGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 152684

RESULT 12
AC020851
LOCUS AC020851 196900 bp DNA linear HTG 15-JUL-2000
```

DEFINITION Mus musculus clone RP21-467L12, WORKING DRAFT SEQUENCE, 33 unordered pieces.

ACCESSION AC020851

VERSION AC020851.2 GI:9211212

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 196900) DOE Joint Genome Institute. Sequencing of Mouse Unpublished

2 (bases 1 to 196900) DOE Joint Genome Institute. Direct Submission

TITLE Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

On Jul 15, 2000 this sequence version replaced gi:6686457.

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

-----

Project Information

Center Project Name: 1437240

Center clone name: RPCI-21\_467L12

-----

Summary Statistics

Consensus quality: 164618 bases at least Q40

Consensus quality: 180887 bases at least Q30

Consensus quality: 183777 bases at least Q20

Estimated insert size: 147000; pulse field gel estimation

Estimated insert size: 193700; sum-of-contrigs estimation

Quality coverage: 4.3 in Q20 bases; pulse field gel estimation

Quality coverage: 4.3 in Q20 bases; sum-of-contrigs estimation

\* NOTE: This is a 'working draft' sequence. It currently consists of 33 contrigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contrigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1541: contrig of 1541 bp in length

1542 1641: gap of unknown length

1642 3492: contrig of 1851 bp in length

3493 3592: gap of unknown length

3593 4738: contrig of 1146 bp in length

4739 4838: gap of unknown length

4839 7249: contrig of 2410 bp in length

7249 7348: gap of unknown length

7349 9179: contrig of 1830 bp in length

9179 9279: gap of unknown length

9279 11494: contrig of 2215 bp in length

11494 11593: gap of unknown length

11594 13607: contrig of 2014 bp in length

13608 13707: gap of unknown length

13708 15947: contrig of 2239 bp in length

15947 16047: gap of unknown length

16047 18392: contrig of 2346 bp in length

18393 18492: gap of unknown length

18493 21231: contrig of 2739 bp in length

21232 21331: gap of unknown length

21332 24562: contrig of 3331 bp in length

24563 24762: gap of unknown length

24763 27140: contrig of 2378 bp in length

27141 27240: gap of unknown length

27241 29950: contrig of 2710 bp in length

29951 30050: gap of unknown length

30051 32245: contrig of 2195 bp in length

32246 32345: gap of unknown length

32346 35688: contrig of 3343 bp in length



ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
REFERENCE  
AUTHORS Welch A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.  
TITLE B7-like molecules and uses thereof  
JOURNAL Patent: WO 0200710-A 11 03-JAN-2002;  
Angen, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1. .895  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
53..865  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD28221.1"  
/db\_xref="GI:19575337"  
/translation="MEGSRDVLAVILVLAQLTAGSSVQIIEGPQNTVLKDSBAHF  
NCTVTHGKLMWLTNOMVLSLTTCGPIITNNRTYASYNSTDSFISLIIHDVQPS  
DSGSVQCSLQNSHGSGFSAFLSDQSGEGPALPTWAILLAVAFSLLLIILVLI  
FCCASRRKESESYONEIRKSANNRTNKADPTKLSKENYGYSSDEAKAAQTAS  
LPPKAEVSLPESKSSLPYQELNKHQPGPATHPRVPSFDIASPQKVRNVTLV"  
BASE COUNT 250 a 223 c 212 g 210 t  
ORIGIN  
Query Match 15.8%; Score 186; DB 6; Length 895;  
Best Local Similarity 71.1%; Pred. No. 1.7e-33;  
Matches 246; Conservative 0; Mismatches 100; Indels 0; Gaps 0;  
QY 24 GAGATGGGGCTGTGATTTCCTCCACGGTTCCTGGGTCTGGTAAATCAAGTCATAGAAGC 83  
Db 83 GNGCTGGTCATCTGGCTCAGTCGACAGCTTCGGATCCAGTATCAGATCATAGAAGT 142  
QY 84 CCCGAGATGAAAGTCTGAGGGCTCCAGGCTCGCTTCAACTGCACCGCTCTCCAG 143  
Db 143 CCTCAGATGTAACAGTCTCAAGGACTCAGAGGCTCACTTCAACTGCACCGCTGACTCAC 202  
QY 144 GCTTGAAGCTCATATGATGGGCTCTCAGTGACATGGTGTGCTAGCGTCAGGCCCATG 203  
Db 203 GCTCGAAGCTTCTCATGFGAGTCTTAACCAATGGTGTGCTGAGTCTCACCCCA 262  
QY 204 GAGCCCATCATCAACCAATGACCGCTTCACTCTCAGAGGTACGACCCAGGGCGGAACCTTC 263  
Db 263 GGACCCATCATCAACCAACCGCTTCACTATGCGATGACGATTAACAGCACTACAGCTTC 322  
QY 264 ACCTCGGAGATGATCATCCAAATGTGGAGCCAGTGATTCGGGAAACATCAGATGCAGC 323  
Db 323 ATCTCGGAGTTGATCATCCATGATGTGACGCCAGTGACTCGGGATCCGTGCAATGCAGC 382  
QY 324 CTCGAGACAGTCGGCTGCATGGATCTGCTTACCTTACCGTCCAAG 369  
Db 383 CTGCGAGACAGCCATGGGTTGGATCTGCTTCTCTCAGTGC 428  
RESULT 15  
BC004806  
LOCUS BC004806 1556 bp mRNA linear ROD 16-APR-2003  
DEFINITION Mus musculus RIKEN cDNA 2010003D20 gene, mRNA (cDNA clone MGC:7960  
IMAGE:3584645), complete cds.  
ACCESSION BC004806  
VERSION BC004806.1 GI:13435932  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1556)  
REFERENCE  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Alesch, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, S.I., Wang, J., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schneitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carrinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalau, D.E.,  
Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 1556)  
Strausberg, R.  
Direct Submission  
Submitted (21-MAR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: amg@bcm.tmc.edu  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louseged, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAC Plate: 10 Row: m Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.  
Location/Qualifiers  
1. .1556  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:7960 IMAGE:3584645"  
/tissue\_type="Mammary tumor. Metallothionien-TGF alpha  
model. 10 month old virgin mouse. Taken by biopsy."  
/clone\_lib="NCI CGAP Mam1"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1. .1556  
/gene="2010003D20Rik"  
/db\_xref="LocusID:72058"  
/db\_xref="MGI:1919308"  
55..867  
/codon\_start=1  
/product="RIKEN cDNA 2010003D20"  
/protein\_id="AAH04806.1"  
/db\_xref="GI:13435933"  
/db\_xref="LocusID:72058"  
/translation="MEGSRDVLAVILVLAQLTAGSSVQIIEGPQNTVLKDSBAHF  
NCTVTHGKLMWLTNOMVLSLTTCGPIITNNRTYASYNSTDSFISLIIHDVQPS  
DSGSVQCSLQNSHGSGFSAFLSDQSGEGPALPTWAILLAVAFSLLLIILVLI  
FCCASRRKESESYONEIRKSANNRTNKADPTKLSKENYGYSSDEAKAAQTAS  
LPPKAEVSLPESKSSLPYQELNKHQPGPATHPRVPSFDIASPQKVRNVTLV"  
CDS

BASE COUNT		447 a	364 c	367 g	378 t
ORIGIN					
Query Match		15.8%; Score 186; DB 10; Length 1556;			
Best Local Similarity		71.1%; Pred No. 1.7e-33;			
Matches 246; Conservative		0; Mismatches 100; Indels 0; Gaps 0;			
Qy	24	GAGATGGGGCTTGTGATTTTCCTCCACGGTTCTGGGTCTGGTAATGAAGTCATAGAAGGC	83		
Db	85	GTGCTGGTTCATCCTGGCTCAGCTGACAGCTCCGGATCCAGTTATCAGATCATAGAAGGT	144		
Qy	84	CCCCAGATGCAACAGTCTTGAAGGCTCCAGGGCTGCTTCAACTGCACCGTCTCCCGAG	143		
Db	145	CCTCAGAAATGPAACAGTCTTAAAGACTCAGAGGCTCATTCAACTGCACCGTGACTCAC	204		
Qy	144	GGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGGTGTGCTAAGCGTCAGGCCCATG	203		
Db	205	GGCTGGAAGCTTCTCATGTGGACTCTTAACCAAAATGGTGTGCTGAGTCTCACCAACCAA	264		
Qy	204	GAGCCCATCATCAACCAATGACGGCTTCACCTCTCAGAGGTAGGACCGGGGGAACTTC	263		
Db	265	GGACCCATCATCAACCAACACCGCTTCACCTATGCCATTACACAGCACTGACAGCTTC	324		
Qy	264	ACCTCGGAGATGATCATCCAAATGTGGAGCCAGTGATTTCGGGGAACATCAGATGCAGC	323		
Db	325	ATCTCGGAGTTGATCATCCATGATGTGCAGCCAGTGACTCGGGATCCGTGCAATGCAGC	384		
Qy	324	CTCCAGAACAGTCGCCTGCATGGATCTGCTTACCTTACCGTCCAAAG	369		
Db	385	CTGCAGAACAGCCATGGGTTTGGATCTGCCCTTCCTCTCAGTGCAAG	430		

Search completed: November 12, 2003, 20:56:28  
Job time : 4559.26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:39:35 ; Search time 347.614 Seconds  
(without alignments)  
9124.610 Million cell updates/sec

Title: US-09-729-264-1  
Perfect score: 1175  
Sequence: 1 ctgtctgccactctgaataa.....gtaatacaactgtatag 1175

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
N_Geneseq_19Jun03.*			
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		
25:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1175	100.0	1175	24	ABK13028
2	1123.4	95.6	1240	24	ABK13030
3	1122.4	95.5	1168	24	ABK13029
4	1067.4	90.8	1139	24	ABK13031
5	779.2	66.3	1392	23	AAS92356
6	381	32.4	1195	24	ABK13032
7	190.4	16.2	401	22	AAT36582
8	186	15.8	754	24	ABK13034

9	186	15.8	895	24	ABK13033	DNA encoding mouse
10	142	12.1	398	25	ABX55016	Bovine EST associa
11	53.4	4.5	277	21	ABX10239	Trinucleotide repe
12	52.8	4.5	7029	23	ABL29756	Drosophila melanog
13	51.8	4.4	389	21	ABX10248	Tri-nucleotide rep
14	51.8	4.4	397	20	ABX89891	Spinocerebellar at
15	51.8	4.4	403	21	ABK10240	Trinucleotide repe
16	49.4	4.2	46954	23	ABK16830	Drosophila melanog
17	48.4	4.1	3217	10	AA91578	Rat androgen recep
18	48.4	4.1	3217	12	AAQ12002	Full-length rat an
19	48.4	4.1	4180	10	AA911773	Rat androgen recep
20	47.8	4.1	381	23	ABV54466	Human prostate exp
21	47.6	4.1	2733	20	ABX84442	Mouse brain CNG-1
22	47.4	4.0	3894	23	ABL03353	Drosophila melanog
23	47.4	4.0	22341	23	ABL03352	Drosophila melanog
24	46.6	4.0	2333	22	AAF75339	Human TGF-beta rec
25	46.4	3.9	1983	23	ABL24483	Drosophila melanog
26	46.4	3.9	4044	23	ABL24482	Drosophila melanog
27	46	3.9	3577	25	ABX16536	Canine cDNA encodi
28	45.8	3.9	2682	23	ABL16097	Drosophila melanog
29	45.8	3.9	8821	23	ABL16096	Drosophila melanog
30	45.6	3.9	1509	23	ABL24571	Drosophila melanog
31	45.6	3.9	3884	23	ABL24570	Drosophila melanog
32	45.4	3.9	3042	23	ABL28445	Drosophila melanog
33	45.4	3.9	3135	23	ABL25653	Drosophila melanog
34	45.4	3.9	5215	23	ABL28444	Drosophila melanog
35	45.4	3.9	5447	23	ABL25652	Drosophila melanog
36	45.2	3.8	5397	23	ABL29757	Drosophila melanog
37	45	3.8	462	23	ABL26941	Drosophila melanog
38	45	3.8	2115	23	ABL26940	Drosophila melanog
39	45	3.8	2260	23	ABL03191	Drosophila melanog
40	45	3.8	3210	23	ABL08977	Drosophila melanog
41	45	3.8	3624	23	ABL03190	Drosophila melanog
42	45	3.8	5975	23	ABL08976	Drosophila melanog
43	45	3.8	6541	23	ABL04172	Drosophila melanog
44	44.8	3.8	254	21	AAZ45758	Polymorphic region
45	44.8	3.8	1024	25	ABZ63769	Toxicologically re

ALIGNMENTS

RESULT 1	
ID	ABK13028 standard; cDNA; 1175 BP.
XX	ABK13028;
AC	ABK13028;
DT	23-APR-2002 (first entry)
XX	DNA encoding human B7-like protein, B7-L_h1.
DE	Human; B7-like protein; B7-L; antiinfertility; gynaecological;
KW	antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
KW	antiinflammatory; dermatological; antipsoriatic; neuroprotective;
KW	antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;
KW	antiasthmatic; nephrotropic; antibacterial; virucide; cancer;
KW	reproductive disorder; graft versus host disease; autoimmune disease;
KW	toxic shock syndrome; allergy; nephropathy; skin disorder;
KW	endocrinopathy; lymphoproliferative disorder; gene; ss.
XX	Homo sapiens.
XX	Key
XX	Location/Qualifiers
FT	27..1175
FT	/*tag= a
FT	/product= "B7-like protein, B7-L_h1"
XX	WO200200710-A2.
PD	03-JAN-2002.
XX	28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.  
 PR 28-NOV-2000; 2000US-0729264.  
 XX (AMGE-) AMGEN INC.  
 XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
 PI WPI; 2002-130881/17.  
 XX P-PSDB; AAU75540.  
 DR New B7-like polypeptides, polynucleotides and their modulators, useful  
 PT for diagnosing, preventing and treating reproductive, immune and  
 PT proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX Claim 1; Fig 1; 135pp; English.  
 PS The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
 CC useful for treating B7-like polypeptide-related disease, disorders or  
 CC conditions including reproductive disorders (e.g. infertility),  
 CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (II) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the coding sequence of human B7-L hi.  
 XX SQ Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 other;

Query Match 100.0%; Score 1175; DB 24; Length 1175;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGTCCTCCCATCTGAATAAAGAGATGGGGCTTGATTTTCCTCCAGGGTCTCGGGT 60  
 DB 1 CTGTCCTCCCATCTGAATAAAGAGATGGGGCTTGATTTTCCTCCAGGGTCTCGGGT 60  
 QY 61 CTGTAATGAAGTCATAGAGGCCCCCAGATGCAACAGTCTTGAAGGGTCCCGAGCTC 120  
 DB 61 CTGTAATGAAGTCATAGAGGCCCCCAGATGCAACAGTCTTGAAGGGTCCCGAGCTC 120  
 QY 121 GCTTCAACTGACCGTCTCCAGGGCTTGAAGCTCATCATGTGGGCTCTCAGTGACATGG 180  
 DB 121 GCTTCAACTGACCGTCTCCAGGGCTTGAAGCTCATCATGTGGGCTCTCAGTGACATGG 180  
 QY 181 TGGTGTAAAGCGTACAGCCCATGAGGCCCATCATCACCATGACCGCTTCACCTCTCAGA 240

Db 181 TGGTGTAAAGCGTACAGGCCCATGAGGCCCATCATCACCATGACCGCTTCACCTCTCAGA 240  
 QY 241 GGTACGACAGGGCGGGAACCTTCCCTCGGAGATGATCATCCAAATGTGAGGCCAGTG 300  
 Db 241 GGTACGACAGGGCGGGAACCTTCCCTCGGAGATGATCATCCAAATGTGAGGCCAGTG 300  
 QY 301 ATTCCGGGAACATCAGATGACGCTCCAGAACAGTCGCTGCATCGATCTGCTTACCTTA 360  
 Db 301 ATTCCGGGAACATCAGATGACGCTCCAGAACAGTCGCTGCATCGATCTGCTTACCTTA 360  
 QY 361 CCGTCCAAAGTTATGGGAGAGCTGTTCAATCCAGTGTAAATCTTGTAGTCGTGAGAATG 420  
 Db 361 CCGTCCAAAGTTATGGGAGAGCTGTTCAATCCAGTGTAAATCTTGTAGTCGTGAGAATG 420  
 QY 421 AACCTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCGGGCTCCGGATATTTCTGGG 480  
 Db 421 AACCTTGTGAAGTTACTTGTCTACCTCAGCTGAGACCGGGCTCCGGATATTTCTGGG 480  
 QY 481 AGCTCGGTCTCTCGTCAAGCCATTCAAGCTATTAATTTGTTCCGAGGCCAGCGACCTTC 540  
 Db 481 AGCTCGGTCTCTCGTCAAGCCATTCAAGCTATTAATTTGTTCCGAGGCCAGCGACCTTC 540  
 QY 541 AAAGTGCAGTGCATCTCTGCTCTGACCCACAGAGCAATGGGACTTTGACTTGCCTGG 600  
 Db 541 AAAGTGCAGTGCATCTCTGCTCTGACCCACAGAGCAATGGGACTTTGACTTGCCTGG 600  
 QY 601 CTACTCTGAAGAGCTGAAAGCCCGCAAGCTCTGCAACTGTAATCTCACTGATTCGGT 660  
 Db 601 CTACTCTGAAGAGCTGAAAGCCCGCAAGCTCTGCAACTGTAATCTCACTGATTCGGT 660  
 QY 661 GTCCCCAAGACACTGGAGGTGGTATTAATATTCAGGTGTATTAATCAAGTTTACCGATT 720  
 Db 661 GTCCCCAAGACACTGGAGGTGGTATTAATATTCAGGTGTATTAATCAAGTTTACCGATT 720  
 QY 721 TAGTGTTCCTATTCCTACTTGGGGGAAAGTGGACTTGGACTAGCAGGCCACCATGCTTC 780  
 Db 721 TAGTGTTCCTATTCCTACTTGGGGGAAAGTGGACTTGGACTAGCAGGCCACCATGCTTC 780  
 QY 781 TGAGCCGACGTGACTCTTACAACTACGCTGCTGCTGCGCGCGCTGTTGTTGGCT 840  
 Db 781 TGAGCCGACGTGACTCTTACAACTACGCTGCTGCTGCGCGCGCTGTTGTTGGCT 840  
 QY 841 GCAACTGCTGCTGCGCTGTTGTTTCTGCTGTAGAGAAAAGAGGATTTCTGATTTCAAT 900  
 Db 841 GCAACTGCTGCTGCGCTGTTGTTTCTGCTGTAGAGAAAAGAGGATTTCTGATTTCAAT 900  
 QY 901 TTCAAAAGAAATCTGAAAAGAGAGACAACAAGAAACTGAGACAGAAAGTGGAAATG 960  
 Db 901 TTCAAAAGAAATCTGAAAAGAGAGACAACAAGAAACTGAGACAGAAAGTGGAAATG 960  
 QY 961 AAAAATCCGGCTACAAATTCAGATGAACAAAAGACACAGACCGCTTCTCTCCCTCCA 1020  
 Db 961 AAAAATCCGGCTACAAATTCAGATGAACAAAAGACACAGACCGCTTCTCTCCCTCCA 1020  
 QY 1021 AATCTGTGAATCAAGTATCTGAAACAAAGAAACACTAGCTGTGGCCCTCTCACCAGC 1080  
 Db 1021 AATCTGTGAATCAAGTATCTGAAACAAAGAAACACTAGCTGTGGCCCTCTCACCAGC 1080  
 QY 1081 GGGCTGATCAACCTGCCACCGCCAGAGTATCCACAGGCTTCTTTAATCTGGCCA 1140  
 Db 1081 GGGCTGATCAACCTGCCACCGCCAGAGTATCCACAGGCTTCTTTAATCTGGCCA 1140  
 QY 1141 GTCTGAGAGGTGAGTAATACAACTGTAGTATAG 1175  
 Db 1141 GTCTGAGAGGTGAGTAATACAACTGTAGTATAG 1175

RESULT 2  
 ABK13030  
 ID ABK13030 standard; cDNA; 1240 BP.  
 XX  
 AC ABK13030;

XX 23-APR-2002 (first entry)  
 XX DNA encoding human B7-like protein, B7-L\_h3.  
 XX Human; B7-like protein; B7-L; antiinfectivity; gynaecological;  
 KW antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;  
 KW antiasthmatic; nephrotropic; antibacterial; virucide; tumor; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT CDS 80..1240  
 FT /tag= a  
 FT /product= "B7-like protein, B7-L\_h3"  
 XX WO200200710-A2.  
 XX 03-JAN-2002.  
 XX 28-JUN-2001; 2001WO-US20719.  
 XX 28-JUN-2000; 2000US-214512P.  
 PR 28-NOV-2000; 2000US-0729264.  
 XX (AMGE-) AMGEN INC.  
 XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
 XX WPI; 2002-130861/17.  
 DR P-PSDB; AAU75542.  
 XX New B7-like polypeptides, polynucleotides and their modulators, useful  
 PT for diagnosing, preventing and treating reproductive, immune and  
 PT proliferative disorders, e.g. cancer and arteriosclerosis -  
 XX Claim 1; Fig 3; 135pp; English.  
 XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are  
 CC useful for treating B7-like polypeptide-related disease, disorders or  
 CC conditions including reproductive disorders (e.g. infertility),  
 CC miscarriage, preterm labour and delivery and endometriosis) and  
 CC proliferative disorders. Antibodies, soluble proteins comprising  
 CC extracellular domains and other regulators of B7-L polypeptides are  
 CC useful for enhancing the immune response to tumours. (I) plays a role in  
 CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions,  
 CC  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the coding sequence of human B7-L\_h3.  
 XX  
 SQ Sequence 1240 BP; 318 A; 319 C; 305 G; 298 T; 0 other;  
 Query Match 95.68; Score 1123.4; DB 24; Length 1240;  
 Best Local Similarity 99.98; Pred. No. 4.2e-302;  
 Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 51 GGTTCGGGTCTGGTAAATGAAGTCATAGAGAGGCCGCCAGAAATCAACAGTCCTGAAGGCG 110  
 DB 116 GGTTCGGGTCTGGTAAATGAAGTCATAGAGAGGCCGCCAGAAATCAACAGTCCTGAAGGCG 175  
 QY 111 TCCAGAGCTCGCTCAACTGACCGCTCCAGAGGCTGGAAGCTCATCATGTGGGCTCTC 170  
 DB 176 TCCAGAGCTCGCTCAACTGACCGCTCCAGAGGCTGGAAGCTCATCATGTGGGCTCTC 235  
 QY 171 AGTGACATGTGTGTCTAAGCGTCAGGCCCATCGAGCCCATCATCAATGACCGCTTC 230  
 DB 236 AGTGACATGTGTGTCTAAGCGTCAGGCCCATCGAGCCCATCATCAATGACCGCTTC 295  
 QY 231 ACCTCTCAGAGGTACGACAGGCGGGAATTCACCTTCGAGAGATGATCATCCCAATGTG 290  
 DB 296 ACCTCTCAGAGGTACGACAGGCGGGAATTCACCTTCGAGAGATGATCATCCCAATGTG 355  
 QY 291 GAGCCCATGTATTCGGGGAAATCAGATGAGCTCCAGAACAGTCCCTCGCATGGATCT 350  
 DB 356 GAGCCCATGTATTCGGGGAAATCAGATGAGCTCCAGAACAGTCCCTCGCATGGATCT 415  
 QY 351 GCTTACCTTACCGTCCCAAGTTATGGGAGAGCTGTTCAATCCAGTGTAAATCTTGTAGTC 410  
 DB 416 GCTTACCTTACCGTCCCAAGTTATGGGAGAGCTGTTCAATCCAGTGTAAATCTTGTAGTC 475  
 QY 411 GCTGAGATGAACCTTGTGAGTTACTTGTCTACCTCAGCTGACCTGGACCCGGCTCCCGAT 470  
 DB 476 GCTGAGATGAACCTTGTGAGTTACTTGTCTACCTCAGCTGACCTGGACCCGGCTCCCGAT 535  
 QY 471 ATTTCTCTGGAGCTCGGTCTCTCGTCAAGCTATTCAGCTATTATTTGTTCCGGAGCCC 530  
 DB 536 ATTTCTCTGGAGCTCGGTCTCTCGTCAAGCTATTCAGCTATTATTTGTTCCGGAGCCC 595  
 QY 531 AGCAGCTTCAAGTGCAGTGAGCATCCTGGCTGTGACCCACAGAGCAATGGACATTG 590  
 DB 596 AGCAGCTTCAAGTGCAGTGAGCATCCTGGCTGTGACCCACAGAGCAATGGACATTG 655  
 QY 591 ACTTGGCTGGTACCTCGAGAGGCTCGAGGCCCGCAAGTCTGCAACTGTAAATCTCACT 650  
 DB 656 ACTTGGCTGGTACCTCGAGAGGCTCGAGGCCCGCAAGTCTGCAACTGTAAATCTCACT 715  
 QY 651 GTGATTCGGTGTCCCAAGACATCGAGGTGGTATTAATATTCAGGTGATTAATCAAGT 710  
 DB 716 GTGATTCGGTGTCCCAAGACATCGAGGTGGTATTAATATTCAGGTGATTAATCAAGT 775  
 QY 711 TTACCGAGTTAGGTTTTTCAATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCAGGC 770  
 DB 776 TTACCGAGTTAGGTTTTTCAATTCCTACTTGGGGCAAGTTGGACTTGGACTAGCAGGC 835  
 QY 771 ACCATGCTTCTGAGCCGAGCTGTACTTCAATAAGCTGTCTGTCTGCTCGCCGCTCGT 830  
 DB 836 ACCATGCTTCTGAGCCGAGCTGTACTTCAATAAGCTGTCTGTCTGCTCGCCGCTCGT 895  
 QY 831 TGTGTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890  
 DB 896 TGTGTGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955  
 QY 891 CGTATTCAATTTCAAAAGAAATCTGAAAAGAGAGAGCAAAAGAAAGAACTGAGACAGAA 950  
 DB 956 CGTATTCAATTTCAAAAGAAATCTGAAAAGAGAGAGCAAAAGAAAGAACTGAGACAGAA 1015

QY 951 AGTGAATGAAACTCCGGCTACAAATTCAGATGAAACAAAAGACACACAGACACCGCTTCT 1010  
 Db 1016 AGTGAATGAAACTCCGGCTACAAATTCAGATGAAACAAAAGACACACAGACACCGCTTCT 1075  
 QY 1011 CTCCTCCCAATCTGTGAATCCAGTGTATCTGACAAAGAAACAGTAGTGTGGCCCT 1070  
 Db 1076 CTCCTCCCAATCTGTGAATCCAGTGTATCTGACAAAGAAACAGTAGTGTGGCCCT 1135  
 QY 1071 COTCAACAGCGGCTGATCAAGTCCACCCAGCCAGCAAGTCAATCCACAGGCTTCTTTT 1130  
 Db 1136 COTCAACAGCGGCTGATCAAGTCCACCCAGCCAGCAAGTCAATCCACAGGCTTCTTTT 1195  
 QY 1131 AATCTGGCAGTCTGAGAGGTCAGTAACTAGTATAG 1175  
 Db 1196 AATCTGGCAGTCTGAGAGGTCAGTAACTAGTATAG 1240

## RESULT 3

ABK13029

ID ABK13029 standard; cDNA; 1168 BP.

XX AC

XX ABK13029;

XX 23-APR-2002 (first entry)

XX

XX DNA encoding human B7-like protein, B7-L\_h2.

XX

XX Human; B7-like protein; B7-L; antiinfertility; gynaecological;

XX antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;

XX antiinflammatory; dermatological; antiposoriatic; neuroprotective;

XX antididiabetic; haemostatic; antithyroid; antitumor; antiallergic;

XX aniaesthatic; nephrotropic; antibacterial; virucide; tumour; cancer;

XX reproductive disorder; graft versus host disease; autoimmune disease;

XX toxic shock syndrome; allergy; nephropathy; skin disorder;

XX endocrinopathy; lymphoproliferative disorder; gene; ss.

XX

XX Homo sapiens.

XX

XX

XX Key Location/Qualifiers

XX CDS 8..1168

XX /\*tag= a

XX /product= "B7-like protein, B7-L\_h2"

XX

XX WO200200710-A2.

XX

XX

XX PD 03-JAN-2002.

XX

XX PF 28-JUN-2001; 2001WO-US20719.

XX

XX PR 28-JUN-2000; 2000US-214512P.

XX

XX PR 28-NOV-2000; 2000US-0729264.

XX

XX PR (AMGE-) AMGEN INC.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC growth and maintenance of cancer cells based on the observation of  
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
 CC polypeptide. Hence, modulators of (I) are useful for the treatment of  
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
 CC in allograft transplantation, graft versus host disease, T-cell  
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L  
 CC molecules are useful for alleviating the symptoms associated with  
 CC diseases involving chronic immune cell dysfunction or to treat  
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
 CC purpura and psoriasis, chronic inflammatory disease such as  
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
 CC are also useful as immunosuppressive agents for bone marrow and organ  
 CC transplantation or to prolong graft survival. B7-L molecules are also  
 CC useful for diagnosis and treatment of diseases involving abnormal cell  
 CC proliferation, including arteriosclerosis and vascular restenosis.  
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic  
 CC shock syndrome or allosensitisation due to blood transfusions, and for  
 CC treatment of allergy, asthma and hypersensitivity reactions, and  
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
 CC lymphoproliferative disorders such as multiple myeloma. The present  
 CC sequence represents the coding sequence of human B7-L\_h2.  
 XX

SQ Sequence 1168 BP; 299 A; 302 C; 288 G; 279 T; 0 other;

Query Match 95.5%; Score 1122.4; DB 24; Length 1168;

Best Local Similarity 99.5%; Pred. No. 7.7e-302;

Matches 1126; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 44 CCTCACGGTTCTGGGTCTGGTAATGAAGTCATGAGAGGCCCCAGAAATGCAACAGTCT 103  
 Db 37 CCCACCCGGTTCTGGGTCTGGTAATGAAGTCATGAGAGGCCCCAGAAATGCAACAGTCT 96  
 QY 104 GAAGGGCTCCACAGGCTCGCTTCACTGCACCGTCTCCAGGGCTGGAGCTCATCATGTG 163  
 Db 97 GAAGGGCTCCACAGGCTCGCTTCACTGCACCGTCTCCAGGGCTGGAGCTCATCATGTG 156  
 QY 164 GGCTCTCAGTGACATGGTGTGTGTAAGCGTCAGGCCCATGAGGCCCATCATCAATGA 223  
 Db 157 GGCTCTCAGTGACATGGTGTGTGTAAGCGTCAGGCCCATGAGGCCCATCATCAATGA 216  
 QY 224 CCGCTTCACTCTCAGAGGTACGACCGGGGGGAACTTACCTCGGAGATGATCATCCA 283  
 Db 217 CCGCTTCACTCTCAGAGGTACGACCGGGGGGAACTTACCTCGGAGATGATCATCCA 276  
 QY 284 CAATGTGGAGGCCAGTGAATTCGGGGAAACATCAGATGAGCGCTCCAGAACAGTCGCTGCA 343  
 Db 277 CAATGTGGAGGCCAGTGAATTCGGGGAAACATCAGATGAGCGCTCCAGAACAGTCGCTGCA 336  
 QY 344 TGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATTTCCAGTGTAAATCT 403  
 Db 337 TGGATCTGCTTACCTTACCGTCCAAAGTTATGGGAGAGCTGTTCATTTCCAGTGTAAATCT 396  
 QY 404 TGTAGTCGCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTTCACTGGAGCCCGGCT 463  
 Db 397 TGTAGTCGCTGAGATGAACCTTGTGAAGTTACTTGTCTACCTTCACTGGAGCCCGGCT 456  
 QY 464 CCGGATATTTCTGGGAGCTCGGTCTCTCGGTGAGCGCATTCAGCTATTATTTTGTTC 523  
 Db 457 CCGGATATTTCTGGGAGCTCGGTCTCTCGGTGAGCGCATTCAGCTATTATTTTGTTC 516  
 QY 524 GAGAGCCAGCGACTTCAAGTGCAGTGCATCTCGGCTCTGACCCACAGAGCAATGG 583  
 Db 517 GAGAGCCAGCGACTTCAAGTGCAGTGCATCTCGGCTCTGACCCACAGAGCAATGG 576  
 QY 584 GACTTTGACTTGGTGGCTACCTTACCTGGAGAGCCCTGAGGCCGCGGAGTCTGCACTGTAAA 643

New B7-like polypeptides, polynucleotides and their modulators, useful for diagnosing, preventing and treating reproductive, immune and proliferative disorders, e.g. cancer and arteriosclerosis -

Claim 1; Fig 2; 135pp; English.

The invention relates to an isolated B7-like (B7-L) polypeptide (I). The polypeptide, polynucleotide encoding it and antibody against (I) are useful for treating B7-like polypeptide-related disease, disorders or conditions including reproductive disorders (e.g. infertility, miscarriage, preterm labour and delivery and endometriosis) and proliferative disorders. Antibodies, soluble proteins comprising extracellular domains and other regulators of B7-L polypeptides are useful for enhancing the immune response to tumours. (I) plays a role in

Db 577 GACTTTGACTTGGTGGCTACTGGAAGAGCCTGAAGGCCGCGAAGTCTGCAACTGTA 636  
 Qy 644 TCTACTGTGATTGGTGTCTCCCAAGACACTGGAGGTGGTATTATATATCCAGGTGATT 703  
 Db 637 TCTCACTGTGATTGGTGTCTCCCAAGACACTGGAGGTGGTATTATATATCCAGGTGATT 696  
 Qy 704 ATCAAGTTTACCGAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 763  
 Db 697 ATCAAGTTTACCGAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 756  
 Qy 764 AGCAGGACCAATGCTTCTGAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823  
 Db 757 AGCAGGACCAATGCTTCTGAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 816  
 Qy 824 CGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 883  
 Db 817 CGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 876  
 Qy 884 AGGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAACTGA 943  
 Db 877 AGGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAGAACTGA 936  
 Qy 944 GACAGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1003  
 Db 937 GACAGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 996  
 Qy 1004 CGCTTCTCTCCCTCCCAATCTGTAATCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1063  
 Db 997 CGCTTCTCTCCCTCCCAATCTGTAATCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1056  
 Qy 1064 TGGCCCTCTCCAGCGGGGTGATCAAGTCCAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1123  
 Db 1057 TGGCCCTCTCCAGCGGGGTGATCAAGTCCAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1116  
 Qy 1124 TTCTTTTAACTGGGCGAGTCTGAGAGGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1175  
 Db 1117 TTCTTTTAACTGGGCGAGTCTGAGAGGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1168  
 RESULT 4  
 ABK13031  
 ID ABK13031 standard; cDNA; 1139 BP.  
 XX  
 AC ABK13031;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE DNA encoding human B7-like protein, B7-L\_h4.  
 XX  
 KW Human; B7-like protein; B7-L; antiinfertility; gynaecological;  
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;  
 KW antidiabetic; haemostatic; antithyroid; antituber; antiallergic;  
 KW antiasthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;  
 KW reproductive disorder; graft versus host disease; autoimmune disease;  
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;  
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..1134  
 CDS /\*tag= a  
 FT /product= "B7-like protein, B7-L\_h4"  
 FT  
 FT  
 FT  
 XX WO200200710-A2.  
 PN  
 XX  
 PD 03-JAN-2002.  
 XX  
 XX 28-JUN-2001; 2001WO-US20719.  
 XX  
 XX 28-JUN-2000; 2000US-214512P.  
 PR  
 XX 28-NOV-2000; 2000US-0729264.  
 PR

(AMGE-) AMGEN INC.

Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

WPI; 2002-130881/17.

New B7-like polypeptides, polynucleotides and their modulators, useful for diagnosing, preventing and treating reproductive, immune and proliferative disorders, e.g. cancer and arteriosclerosis -

Claim 1; Fig 4; 135pp; English.

The invention relates to an isolated B7-like (B7-L) polypeptide (I). The polypeptide, polynucleotide encoding it and antibody against (I) are useful for treating B7-like polypeptide-related disease, disorders or conditions including reproductive disorders (e.g. infertility, miscarriage, preterm labour and delivery and endometriosis) and proliferative disorders. Antibodies, soluble proteins comprising extracellular domains and other regulators of B7-L polypeptides are useful for enhancing the immune response to tumours. (I) plays a role in growth and maintenance of cancer cells based on the observation of seminal vesicle hyperplasia in transgenic mice overexpressing B7-L polypeptide. Hence modulators of (I) are useful for the treatment of cancer including seminal vesicle cancer, lung, brain, breast, ovarian, testicular cancer and cancers of haematopoietic system. B7-L polypeptide pathway can be manipulated to regulate cytotoxic T-lymphocyte response in allograft transplantation, graft versus host disease, T-cell dependent B-cell mediated diseases and autoimmune diseases. B7-L molecules are useful for alleviating the symptoms associated with diseases involving chronic immune cell dysfunction or to treat autoimmune diseases such as systemic lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes, immune thrombocytopenic purpura and psoriasis, chronic inflammatory disease such as inflammatory bowel disease (Crohn's disease and ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They are also useful as immunosuppressive agents for bone marrow and organ transplantation or to prolong graft survival. B7-L molecules are also useful for diagnosis and treatment of diseases involving abnormal cell proliferation, including arteriosclerosis and vascular restenosis. Antagonists of B7-L polypeptides are useful for alleviation of toxic shock syndrome or allosensitisation due to blood transfusions, and for treatment of allergy, asthma and hypersensitivity reactions, and for nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and lymphoproliferative disorders such as multiple myeloma. The present sequence represents the coding sequence of human B7-L\_h4.

Sequence 1139 BP; 290 A; 300 C; 283 G; 266 T; 0 other;

Query Match 90.8%; Score 1067.4; DB 24; Length 1139;  
 Best local Similarity 97.5%; Pred. No. 1.5e-286;  
 Matches 1103; Conservative 0; Mismatches 6; Indels 22; Gaps 1;

Qy	44	CTCTCCAGGTCCTGGTCTGGTAAATGAAGTCATAGAGGCCGCCAGAGTCAACAGTCCT	103
Db	30	CCACCCCGTCTGGTCTGGTAAATGAAGTCATAGAGGCCGCCAGAGTCAACAGTCCT	89
Qy	104	GAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTTCCCGGGGTGGAAGCTCATCATGTG	163
Db	90	GAAGGGCTCCAGGCTCGCTTCAACTGCACCGCTTCCCGGGGTGGAAGCTCATCATGTG	149
Qy	164	GGCTCTCAGTGCATGGTGGTGTAGAGCGTCAGGCCCATGGACCCATATATCAATGA	223
Db	150	GGCTCTCAGTGCATGGTGGTGTAGAGCGTCAGGCCCATGGACCCATATATCAATGA	209
Qy	224	CCGCTTCACCTCTCAGAGGTACACACAGCGCGGGAACCTTCACTCCGAGATGATATCA	283
Db	210	CCGCTTCACCTCTCAGAGGTACACACAGCGCGGGAACCTTCACTCCGAGATGATATCA	269
Qy	284	CAATGGAGCCCGAGTGTTCGGGGGAACATCAGATGCAGCCTCCAGAACAGTCGCTGCA	343

Db 270 CAATGTGGACCCAGTGAATTCGGGAAACATCAGATGCAGGCTCCAGAACAGTCGCCCTGCA 329  
QY 344 TGGATCTGCTTACCTTACCGTCAAGTTATGGAGAGCTGTTTCAATCCCAAGTGTAAATCT 403  
Db 330 TGGATCTGCTTACCTTACCGTCAAGTTATGGAGAGCTGTTTCAATCCCAAGTGTAAATCT 389  
QY 404 TGTAGTCGCTGAGAACTTGTGAAGTTACTTGTCTACCTTCACTGGACCCGGCT 463  
Db 390 TGTAGTCGCTGAGAACTTGTGAAGTTACTTGTCTACCTTCACTGGACCCGGCT 449  
QY 464 CCCGATATTCCTGGAGCTCGGTCTCTGGTCAAGCATTCAGCTATTAATTTGTTCC 523  
Db 450 CCCGATATTCCTGGAGCTCGGTCTCTGGTCAAGCATTCAGCTATTAATTTGTTCC 509  
QY 524 GGAGCCAGGACCTTCAAGTGCAGTGAGCATCCTGGCTCTGACCCCAAGCAATGG 583  
Db 510 GGAGCCAGGACCTTCAAGTGCAGTGAGCATCCTGGCTCTGACCCCAAGCAATGG 569  
QY 584 GACTTTGACTTGGTGTACCTGGAGAGCCTGAGGCCGCGCAAGTCTCAACTGTAAA 643  
Db 570 GACTTTGACTTGGTGTACCTGGAGAGCCTGAGGCCGCGCAAGTCTCAACTGTAAA 629  
QY 644 TCTCACTGTGATCGGTGTCCCAAGACACTGGAGGTGGTATTAATATCCAGGTGATT 703  
Db 630 TCTCACTGTGATCGGTGTCCCAAGACACTGGAGGTGGTATTAATATCCAGGTGATT 689  
QY 704 ATCAAGTTACCGAGTTTGGTTCATTCGCTTACTTGGGGCAAGTTGGACTTGGACT 763  
Db 690 ATCAAGTTTACCGAGTTTGGTTCATTCGCTTACTTGGGGCAAGTTGGACTTGGACT 749  
QY 764 AGCAGGACCATGCTTCTGACGCGAGCTGCTTCTTACAAATACGCTGCTGCTGCCG 823  
Db 750 AGCAGGACCATGCTTCTGACGCGAGCTGCTTCTTACAAATACGCTGCTGCTGCCG 809  
QY 824 CCGTGTGTGTGGTGTCAACTGCTGCTCGGCTGTTGTTCTGCTGTAGAGAAAAG 883  
Db 810 CCGTGTGTGTGGTGTCAACTGCTGCTCGGCTGTTGTTCTGCTGTAGAGAAAAG 869  
QY 884 AGGATTCGTATTCAATTTCAAAGAAATCTGAAAGAGAGACAAACAAAGAACTGA 943  
Db 870 AG-----GAAATCTGAAAAGAGAGACAAACAAAGAACTGA 907  
QY 944 GACAGAAAGTGAATGAAATCCGGCTACATTCAGATGAAACAAAGACACAGAC 1003  
Db 908 GACAGAAAGTGAATGAAATCCGGCTACATTCAGATGAAACAAAGACACAGAC 967  
QY 1004 CGTTTCTCTCCCAATCTCTGTAATCCAGTGATCCTTGAAACAAAGAAACAGTAGCTG 1063  
Db 968 CGTTTCTCTCCCAATCTCTGTAATCCAGTGATCCTTGAAACAAAGAAACAGTAGCTG 1027  
QY 1064 TGGCCCTCTCACAGCGGCTGATCAACGCTCCACCCAGCCAGCAAGTCATCCAGGC 1123  
Db 1028 TGGCCCTCTCACAGCGGCTGATCAACGCTCCACCCAGCCAGCAAGTCATCCAGGC 1087  
QY 1124 TTCTTTTAACTCTGGCCAGTCTGAGAGGTCAGTAATACAACCTGTAGTATA 1174  
Db 1088 TTCTTTTAACTCTGGCCAGTCTGAGAGGTCAGTAATACAACCTGTAGTATA 1138

RESULT 5

AAS92356  
ID AAS92356 standard; cDNA; 1392 BP.  
XX  
AC AAS92356;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #28160.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.  
PN WO200175067-A2.  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG28169.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX  
PS Claim 1; SEQ ID No 28160; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1392 BP; 338 A; 361 C; 341 G; 352 T; 0 other;

Query Match 66.3%; Score 779.2; DB 23; Length 1392;  
Best Local Similarity 87.3%; Pred. No. 2.9e-206;  
Matches 913; Conservative 0; Mismatches 23; Indels 110; Gaps 1;  
QY 159 ATGTGGGCTCTCAGTGACATGCTGCTTAAGCGTCAGGCCCATGAGCCCATCATCACC 218  
Db 1 ARGTTGGCTCTCAGTGACATGCTGCTTAAGCGTCAGGCCCATGAGCCCATCATCACC 60  
QY 219 AATGACCGCTTCCCTCTCAGAGGTACGACACAGGCGGGAACCTTCACTCGGAGATGATC 278  
Db 61 AATGACCGCTTCCCTCTCAGAGGTACGACACAGGCGGGAACCTTCACTCGGAGATGATC 120  
QY 279 ATCCCAATGTGGAGCCCATGATTCGGGGAACATCAGATGAGCTTCAGAACAGTCCG 338  
Db 121 ATCCCAATGTGGAGCCCATGATTCGGGGAACATCAGATGAGCTTCAGAACAGTCCG 180  
QY 339 CTGCATGATCTGCTTACCTTACCTTCAAGTTATGGAGAGCTGTTCAATCCAGTGT 398  
Db 181 CTGCATGATCTGCTTACCTTACCTTCAAGTTATGGAGAGCTGTTCAATCCAGTGT 240  
QY 399 AATCTGTAGTCGCTGAGAACTGCTGTGAAGTTACTTGTCTACCTTCACACTGGACC 458  
Db 241 AATCTGTAGTCGCTGAGAACTGCTGTGAAGTTACTTGTCTACCTTCACACTGGACC 300







PD 03-JAN-2002.  
XX 28-JUN-2001; 2001WO-US20719.  
XX 28-JUN-2000; 2000US-2145112P.  
XX 28-NOV-2000; 2000US-0729264.  
XX (AMGE-) AMGEN INC.  
XX Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;  
XX WPI; 2002-130881/17.  
XX P-PSDB; AAU75545.  
XX New B7-like polypeptides, polynucleotides and their modulators. useful  
XX PT for diagnosing, preventing and treating reproductive, immune and  
XX PT proliferative disorders, e.g. cancer and arteriosclerosis -  
XX Claim 1; Fig 6; 135pp; English.  
XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).  
XX The polypeptide, polynucleotide encoding it and antibody against (I) are  
XX useful for treating B7-like polypeptide-related disease disorders or  
XX conditions including reproductive disorders (e.g. infertility, and  
XX miscarriage, preterm labour and delivery and endometriosis) and  
XX proliferative disorders. Antibodies, soluble proteins comprising  
XX extracellular domains and other regulators of B7-L polypeptides are  
XX useful for enhancing the immune response to tumours. (I) plays a role in  
XX growth and maintenance of cancer cells based on the observation of  
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L  
XX polypeptide. Hence modulators of (I) are useful for the treatment of  
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,  
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide  
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response  
XX in allograft transplantation, graft versus host disease, T-cell  
XX dependent B-cell mediated diseases and autoimmune diseases. B7-L  
XX molecules are useful for alleviating the symptoms associated with  
XX diseases involving chronic immune cell dysfunction or to treat  
XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid  
XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic  
XX purpura and psoriasis, chronic inflammatory disease such as  
XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),  
XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They  
XX are also useful as immunosuppressive agents for bone marrow and organ  
XX transplantation or to prolong graft survival. B7-L molecules are also  
XX useful for diagnosis and treatment of diseases involving abnormal cell  
XX proliferation, including arteriosclerosis and vascular restenosis.  
XX Antagonists of B7-L polypeptides are useful for alleviation of toxic  
XX shock syndrome or allo sensitisation due to blood transfusions, and for  
XX treatment of allergy, asthma and hypersensitivity reactions.  
XX nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and  
XX pemphigoid), endocrinopathies (Grave's disease), various pneumopathies  
XX (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,  
XX lymphocytopenias, Guillain-Barre syndrome and myasthenia gravis, and  
XX thromboproliferative disorders such as multiple myeloma. The present  
XX sequence represents the coding sequence of mouse B7-L<sub>m2</sub>.  
XX Sequence 895 BP; 250 A; 223 C; 212 G; 210 T; 0 other;  
XX Query Match 15.8%; Score 186; DB 24; Length 895;  
XX Best Local Similarity 71.1%; Pred. No. 2.9e-41;  
XX Matches 246; Conservative 0; Mismatches 100; Indels 0; Gaps 0;  
XX 24 GAGATGGGGCTGTGATTTCTCCACGGTCTCGGTTCGTAATGAAGTCATAGAAGGC 83  
DB 83 GTGCTGGTATCTGCTGAGTGCAGCTTCGGATCCGATTCAGTATCATAGAAGT 142  
XX 84 CCCAGAAATGCAAGTCTGAAGGGCTCCAGGCTCGCTTCAATGACCGTCTCCAG 143  
DB 143 CCTCAGAAATGCAAGTCTGAAGGGCTCAGAGGCTCACTTCAATGACCGTCACTC 202  
XX 144 GCGTGAAGTCTCATATGTTGGCTCTCAGTGCATGTTGGTGGTCTAAGCGTCAGGCCCATG 203

203 GGCTGGAAGCTTCTCATGTGGACTCTTAACCAAAATGGTGGTCTGAGTCTCACCACCCAA 264  
204 GAGCCCATCATCAACCAATGACCGCTTACCTCTCAGAGGTAGCAGCAGGGGGAACTTC 263  
263 GGACCCATCATCACCACCAACACCGCTTACCTATGCCAGTTTACACAGCACTGACAGCTTC 322  
264 ACTCTGGAGATGATCATCCACAATGTGGAGCCAGTGTTCGGGGAACATCAGATGTCAGC 323  
323 ATCTCGGAGTTCATCATCCATGATGTCAGCCCACTGACTCGGATCCGTGCAATGTCAGC 382  
324 CTCCAGAACAGTCCGCTGTCATGATGTCGTTACCTTACCTTACGTTCCAAAG 369  
383 CTGCAGAACAGCCATGGGTTGGATCTGCTCTCTCTCAGTCCAAAG 428

RESULT 10  
ABX55016  
ID ABX55016 standard; cDNA; 398 BP.  
XX AC ABX55016;  
XX XX  
XX 26-FEB-2003 (first entry)  
XX DE Bovine EST associated with lactation/muscle/fat deposition #4945.  
XX XX  
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
XX muscle deposition; fat deposition; genome mapping; gene identification;  
XX gene analysis; cattle breeding.  
XX OS Bos Taurus.  
XX XX  
XX US2002137160-A1.  
XX XX  
XX 26-SEP-2002.  
XX PF 26-OCT-2001; 2001US-0983965.  
XX XX  
XX 17-DEC-1998; 98US-113678P.  
XX PR 15-DEC-1999; 99US-0469231.  
XX XX  
XX (BYAT/) BYATT J C.  
XX (MATH/) MATHIALAGAN N.  
XX (TAON/) TAO N.  
XX (WARR/) WARREN W C.  
XX XX  
XX Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX WPI; 2003-102386/09.  
XX PT Purified nucleic acid molecules, useful for genome mapping, gene  
XX identification and analysis, cattle breeding or preparation of  
XX constructs for cattle gene expression and genetically improved cattle -  
XX Claim 2; SEQ ID No 4945; 38pp; English.  
XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived  
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a  
XX second nucleic acid molecule comprising any of 5912 nucleotide  
XX sequences, appearing as ABX50072-ABX5983, or complements of them.  
XX Also included are; (1) a transformed cell having a nucleic acid  
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
XX translated sequence that functions in the cell to cause termination of  
XX transcription and addition of polyadenylated ribonucleotides to a 3' end  
XX of the mRNA molecule; and (2) determining a level or pattern of a  
XX molecule in a bovine cell or tissue comprising: (a) incubating a marker  
XX nucleic acid (comprising any of the 5912 nucleic acid sequences or its  
XX complement or fragment) with a complementary nucleic acid molecule  
XX obtained from the bovine cell or tissue, where hybridisation between the  
XX marker nucleic acid and the complementary nucleic acid permits the  
XX detection of the molecule; and (b) detecting the level or pattern of the  
XX complementary nucleic acid, where the detection of the complementary  
XX nucleic acid is predictive of the level or pattern of the molecule.







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 16:41:25 ; Search time 2816.33 seconds  
(without alignments)  
10140.072 Million cell updates/sec

Title: US-09-729-264-1  
Perfect score: 1175  
Sequence: 1 cftgtgccactctgaataa.....gtaatacaactgtagtatag 1175

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_esti.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pin.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	786.2	66.9	1201	13 BX399881	BX399881 BX399881
2	543.6	46.3	941	13 BX403420	BX403420 BX403420
3	509.2	43.3	917	13 BX370118	BX370118 BX370118
4	498.2	42.4	1201	13 BX358190	BX358190 BX358190

5	456	38.8	970	13 BX370209	BX370209 BX370209
6	328.6	28.0	1510	11 AK086973	AK086973 Mus muscu
7	308.6	26.3	784	12 BG863468	BG863468 602796941
8	279	23.7	794	10 BF784177	BF784177 602108039
9	267.4	22.8	725	10 BG740428	BG740428 602633817
10	266.4	22.7	1148	13 BX370210	BX370210 BX370210
11	259.4	22.1	631	10 BB663870	BB663870 BB663870
12	247.8	21.1	527	10 BE032610	BE032610 132035 WA
13	226.2	19.3	1201	13 BX358189	BX358189 BX358189
14	209	17.8	556	4 BX511978	BX511978 RZPD Mus
15	199	16.9	517	10 BE476432	BE476432 159670 BA
16	191.6	16.3	524	9 AV601192	AV601192 AV601192
17	190.8	16.2	622	29 AG069679	AG069679 Pan trogl
18	190.8	16.2	677	29 AG107877	AG107877 Pan trogl
19	186	15.8	471	9 AA265274	AA265274 mx91d09.r
20	186	15.8	611	10 AW990468	AW990468 uf37405.y
21	186	15.8	691	14 BY708046	BY708046 BY708046
22	186	15.8	931	10 EG173684	EG173684 602336823
23	186	15.8	1552	11 AK080860	AK080860 Mus muscu
24	186	15.8	1552	11 AK078934	AK078934 Mus muscu
25	185	15.7	486	9 AI425363	AI425363 mx91d09.y
26	184.4	15.7	663	10 BE376590	BE376590 601226419
27	182.8	15.6	650	10 BB625217	BB625217 BB625217
28	168.2	14.3	600	28 AZ379623	AZ379623 LM0134G22
29	155.2	13.2	388	14 CB805443	CB805443 ANGNNUC.C
30	153.2	13.0	216	10 BG206666	BG206666 RST26117
31	153.2	13.0	556	9 AI788300	AI788300 uk36d01.y
32	151.6	12.9	530	9 AI790785	AI790785 uk28a12.y
33	145.8	12.4	423	10 BB846577	BB846577 BB846577
34	143.8	12.2	291	28 AZ411779	AZ411779 LM0185B04
35	134	11.4	426	13 BY033616	BY033616 BY033616
36	130.6	11.1	351	13 BY100408	BY100408 BY100408
37	128.8	11.0	471	10 BF040046	BF040046 BF250004B
38	127.4	10.8	394	10 BB846133	BB846133 BB846133
39	124	10.6	685	29 AG142221	AG142221 Pan trogl
40	122.6	10.4	346	13 BY143304	BY143304 BY143304
41	121.6	10.3	369	13 BY036814	BY036814 BY036814
42	110	9.4	785	12 BI452873	BI452873 603170211
43	102.8	8.7	889	12 BI454276	BI454276 603171509
44	100	8.5	185	14 N47851	N47851 yw95h05.r1
45	100	8.5	234	14 N93995	N93995 za66f09.r1

## ALIGNMENTS

RESULT 1  
BX399881  
LOCUS  
DEFINITION BX399881 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1085YA24 5-PRIME, mRNA sequence.  
ACCESSION BX399881  
VERSION BX399881  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10299.f, and  
it belongs to a clone representative of this cluster. For more  
information about this cluster and the virtual cDNA, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0D1085BA12QPI&cluster=10299.f. Contact :  
Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS001085BA12QPI.  
Location/Qualifiers  
1. 1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS001085YA24"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 290 a 285 c 301 g 297 t 28 others  
ORIGIN  
Query Match 66.9%; Score 786.2; DB 13; Length 1201;  
Best Local Similarity 93.4%; Pred. No. 4.9e-133;  
Matches 838; Conservative 21; Mismatches 34; Indels 4; Gaps 4;  
QY 52 GTTCTGGGCTGTAATGAAGTCATAGAGAGGCCCCAGAAATCAACAGTCTGAAGGGCT 111  
DB 272 GTTCTGGGCTGTAATGAAGTCATAGAGAGGCCCCAGAAATCAACAGTCTGAAGGGCT 331  
QY 112 CCCAGGCTCGCTTCAACTGCACCGCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 171  
DB 332 CCCAGGCTCGCTTCAACTGCACCGCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 391  
QY 172 GTGACATGTGTGCTTAAGCTGACGCCCATGAGGCCCATCATCATCAACATGACCGCTTCA 231  
DB 392 GTGACATGTGTGCTTAAGCTGACGCCCATGAGGCCCATCATCATCAACATGACCGCTTCA 451  
QY 232 CCTCTCAGAGGTACGACACCGGCGGAACTTCACTCGGAGATGATCATCAATGTGG 291  
DB 452 CCTCTCAGAGGTACGACACCGGCGGAACTTCACTCGGAGATGATCATCAATGTGG 511  
QY 292 AGCCGAGTATTCGGGAGACATGAGTGCAGCTCCAGACAGTGCCTGCATGATCTG 351  
DB 512 AGCCGAGTATTCGGGAGACATGAGTGCAGCTCCAGACAGTGCCTGCATGATCTG 571  
QY 352 CTTCACCTTACCGTCCAAAGTTATGGAGAGCTGTTCACTCCAGTGTAACTTGTAGTCG 411  
DB 572 CTTCACCTTACCGTCCAAAGTTATGGAGAGCTGTTCACTCCAGTGTAACTTGTAGTCG 631  
QY 412 CTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCTCACTGGACCCGGCTCCCGGATA 471  
DB 632 CTGAGATGAACCTTGTGAAGTTACTTGTCTACCTCTCACTGGACCCGGCTCCCGGATA 691  
QY 472 TTTCTCTGGAGCTCGCTCTCTGTCAGCCATTCAGCTATTATTTTGTTCGGAGCCCA 531  
DB 692 TTTCTCTGGAGCTCGCTCTCTGTCAGCCATTCAGCTATTATTTTGTTCGGAGCCCA 751  
QY 532 GCGACCTTCAAGTCAAGTGAAGTCTGCTGCTGACCCCAAGAGCAATGGGACTTTGA 591  
DB 752 GCGACCTTCAAGTCAAGTGAAGTCTGCTGCTGACCCCAAGAGCAATGGGACTTTGA 811  
QY 592 CTTGCGTGGCTACTTGAAGAGCTGAGCCCGCAAGTCTGCACTGTAAATCTCACATG 651  
DB 812 CTTGCGTGGCTACTTGAAGAGCTGAGCCCGCAAGTCTGCACTGTAAATCTCACATG 871  
QY 652 TGATTCGGTGTCCCAAGACATCTGAGGTGGTATTAATATTCAGGTGTATTATCAAGTT 711  
DB 872 TGATTCGGTGT - CCCAAGACATCTGAGGTGGTATTAATATTCAGGTGTATTATCAAG -T 929  
QY 712 TACCGAGTTAGGTTTTTCACTTGGGCGCAAGTTGGACTTGGACTAGCAGGCA 771  
DB 930 TACCGAGTTAGGTTTTTCACTTGGGCGCAAGTTGGACTTGGACTAGCAGGCA 989  
QY 772 CCAATGCTTCTGAGCGCGAGCTGTACTCTTACAATACGCTGTGCTGCTCGCGCGCTGTT 831  
DB 990 CCAATGCTTCTGAGCGCGAGCTGTACTCTAATATATGCTGCTGCTCGCGCGCTGTT 1049

QY 832 GTTGTGGCTGCAACTGCTGCTGCCGTGTTGTTTCTGCTGTAGAGAGAGGATTTTC 891  
DB 1050 GATGKTGCTGCAACT -CTGCKRCGGTAKTGTGTTCTTCTGTAGAA -RAAABARGGTTYS 1107  
QY 892 GATTTCAATTTCAAAAGAAATCTGAAAAAGAGAGAGCAAAACAAAGAACTGAGACAG 948  
DB 1108 GTATTCATTTCAARAATATCTGAAAGAGAGAGCAAAACAAAGAAACGAGACAGMGMG 1164  
RESULT 2  
BX403420 941 bp mRNA linear EST 13-MAY-2003  
LOCUS BX403420 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS001085YA24 5-PRIME, mRNA sequence.  
ACCESSION BX403420  
VERSION BX403420.1 GI:30607302  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 941)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
10299.f, and it belongs to a clone representative of this cluster.  
For more information about this cluster and the virtual cDNA, see  
http://www.genoscope.cns.fr/  
cgl-bin/cluster.cgi?seq=CS5AA017ZE02RM1&cluster=10299.f. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS5AA017ZE02RM1.  
Location/Qualifiers  
1. 941  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS001085YA24"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 214 a 241 c 255 g 227 t 4 others  
ORIGIN  
Query Match 45.3%; Score 543.6; DB 13; Length 941;  
Best Local Similarity 99.0%; Pred. No. 6e-89;  
Matches 568; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 52 GTTCTGGGCTCGTGAATGAAGTCATAGAGAGGCCCCAGAAATCAACAGTCTGAAGGGCT 111  
DB 286 GTTCTGGGCTCGTGAATGAAGTCATAGAGAGGCCCCAGAAATCAACAGTCTGAAGGGCT 345  
QY 112 CCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 171  
DB 346 CCCAGGCTCGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCTCA 405  
QY 172 GTGACATGTGTGCTTAAGCTGACGCCCATGAGGCCCATCATCAACATGACCGCTTCA 231  
DB 406 GTGACATGTGTGCTTAAGCTGACGCCCATGAGGCCCATCATCAACATGACCGCTTCA 465  
QY 232 CCTCTCAGAGGTACGACACCGGCGGAACTTCACTCGGAGATGATCATCAATGTGG 291  
DB 466 CCTCTCAGAGGTACGACACCGGCGGAACTTCACTCGGAGATGATCATCAATGTGG 525

292 AGCCAGTATTCGGGGAACATCAGATGAGCTCCAGAACAGTGCCTGCGATGATCTG 351  
 526 AGCCAGTATTCGGGGAACATCAGATGAGCTCCAGAACAGTGCCTGCGATGATCTG 585  
 352 CTTACCTTACCGTCCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAAATCTTTAGTGG 411  
 586 CTTACCTTACCGTCCCAAGTTATGGGAGAGCTGTTCAATCCCAAGTGTAAATCTTTAGTGG 645  
 412 CTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCCCTCACATGGAGCCCGGCTCCCGGATA 471  
 646 CTGAGAAATGAACCTTGTGAAGTTACTTGTCTACCCCTCACATGGAGCCCGGCTCCCGGATA 705  
 472 TTTCTGGGAGCTCGCTCTCTCGTCAGCAATCAAGCTATTAATTTG-TTCGGAGGCC 530  
 706 TTTCTGGGAGCTCGCTCTCTCGTCAGCAATCAAGCTATTAATTTGTTTCGGAGGCC 765  
 531 AGCGAATCTCAAGTGCAGTGCATCTCTGCTGACCCCTGACCCCAAGCAATGGACTTTG 590  
 766 AGCGAATCTCAAGTGCAGTGCATCTCTGCTGACCCCTGACCCCAAGCAATGGACTTTG 824  
 591 ACTTGGTGTCTACCTGGAGAGCCTGAAGGCC 624  
 825 ACTTGGTGTCTACCTGGAGAGCCTGAAGGCC 858

## RESULT 3

BX370118

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2250.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAG037ZE05\_CS03493\_2&amp;cluster=2250.r.

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAG037ZE05\_CS03493\_2.

Location/Qualifiers

1. .917

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1076YN22"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

523; Conservative

0; Mismatches

3; Indels

2; Gaps

1;

917

917

917

917

917

917

917

917

917

917

917

917

917

9 ACTGTGATTCGGTGTCCCCA--GAACTGGAGGTGGTATTATATTCAGGTGATTATCA 66  
 708 AGTTTACCGAGTTTGTAGGTTTTCATTCGCTACTTGGGGGAAAAGTTGGACTTGAAGTAGCA 767  
 67 AGTTTACCGAGTTTGTAGGTTTTCATTCGCTACTTGGGGGAAAAGTTGGACTTGAAGTAGCA 126  
 768 GGCACCATGCTTCGAGCGGAGCGTGTACTCTTACAATAACGCTGCTGCTGCGCGCGT 827  
 127 GGCACCATGCTTCGAGCGGAGCGTGTACTCTTACAATAACGCTGCTGCTGCGCGCGT 186  
 828 CGTTGTTGTGGCTCAAACTGCTGCTGCGTGTGTTTCTGCTGTAGAGAAAAGAGGA 887  
 187 CGTTGTTGTGGCTCAAACTGCTGCTGCGTGTGTTTCTGCTGTAGAGAAAAGAGGA 246  
 888 TTTGCTGTTCAATTTCAAAAGAAATCTGAAAAGAGAGAGCAAAACAAAGAAATCTGAGACA 947  
 247 TTTGCTGTTCAATTTCAAAAGAAATCTGAAAAGAGAGAGCAAAACAAAGAAATCTGAGACA 306  
 948 GAAAGTGGAAATGAAAACCTCCGGCTACAATTCAGATGAACAAAAGACACACAGACCCGCT 1007  
 307 GAAAGTGGAAATGAAAACCTCCGGCTACAATTCAGATGAACAAAAGACACACAGACCCGCT 366  
 1008 TCTCTCCCTCCCAATCTGTGAATCCAGTATCTGAAAGAGAGAGAGCAAAACAAAGAAATCTGAGACA 1067  
 367 TCTCTCCCTCCCAATCTGTGAATCCAGTATCTGAAAGAGAGAGAGCAAAACAAAGAAATCTGAGACA 426  
 1068 CTTCTCTACCAAGCGGCTGATCAACAGTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCT 1127  
 427 CTTCTCTACCAAGCGGCTGATCAACAGTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCT 486  
 1128 TTTAATCTGGCCAGTCTGAGAAGCTGATGAATCACTGATGATAG 1175  
 487 TTTAATCTGGCCAGTCTGAGAAGCTGATGAATCACTGATGATAG 534

## RESULT 4

BX358190

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10299.f

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0D1035AB01QPI.

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1035YCO1"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

Query Match

Best Local Similarity

Matches

523; Conservative

0; Mismatches

3; Indels

2; Gaps

1;

917

917

917

917

```

BASE COUNT      341 a      245 c      245 g      319 t      51 others
ORIGIN
Query Match      42.4%; Score 498.2; DB 13; Length 1201;
Best Local Similarity 95.7%; Pred. No. 1e-80;
Matches 512; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 641 AAATCTCACTGTGATTCGGTGTCCCAAGACACTGAGGTGGTATTATATATCCAGGTGT 700
Db 27 AAAAAAGCAGCGTGGTACCGGTCCGGAATTCGGGATGGTGGTATTAATATCCAGGTGT 86
QY 701 ATTATCAAGTTTACCGAGTTTAGGTTTTCATTGCTACTTGGGCAAGTTGACACTGG 760
Db 87 ATTATCAAGTTTACCGAGTTTAGGTTTTCATTGCTACTTGGGCAAGTTGACACTGG 146
QY 761 ACTAGAGCAGCATGCTTCTGACCGGACGCTGCTTCAATACACCTGCTGCTGCTG 820
Db 147 ACTAGAGCAGCATGCTTCTGACCGGACGCTGCTTCAATACACCTGCTGCTGCTG 206
QY 821 CCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
Db 207 CCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 266
QY 881 AAGAGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAGAGAGAGACAAAGAAAC 940
Db 267 AAGAGATTTCTGATTTCAATTTCAAAAGAAATCTGAAAGAGAGAGACAAAGAAAC 326
QY 941 TGAGACAGAAAGTGGAAATGAAATCTCCGCTTCAATTCAGATGAACAAAGACACAGA 1000
Db 327 TGAGACAGAAAGTGGAAATGAAATCTCCGCTTCAATTCAGATGAACAAAGACACAGA 386
QY 1001 CACCGCTTCTCCCTCCCAATTCCTGTAATCCAGTATGCTGAAACAAAGAAACAGTAG 1060
Db 387 CACCGCTTCTCCCTCCCAATTCCTGTAATCCAGTATGCTGAAACAAAGAAACAGTAG 446
QY 1061 CTGTGCGCTTCTCCACCGGCGTATCAAGTCCACCCGCGCAGAGAGTATCCACA 1120
Db 447 CTGTGCGCTTCTCCACCGGCGTATCAAGTCCACCCGCGCAGAGAGTATCCACA 506
QY 1121 GGCTTCTTTTAACTCTGGCAGCTCTGAGAAAGTTCAGTATATCAACTGTAGTATAG 1175
Db 507 GGCTTCTTTTAACTCTGGCAGCTCTGAGAAAGTTCAGTATATCAACTGTAGTATAG 561

RESULT 5
BX370209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
LOCUS clone CS0D1085YA24 5-PRIME, mRNA sequence.
DEFINITION
ACCESSION BX370209.1 GI:30451880
VERSION BX370209
KEYWORDS ESI.
SOURCE Homo sapiens (human)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
COMMENT 1 (bases 1 to 970)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10299.f, and
it belongs to a clone representative of this cluster. For more
information about this cluster and the virtual cDNA, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG039ZE05_C503685_1scluster=10299.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BAG039ZE05_C503685_1.

FEATURES
source
Location/Qualifiers
1..970
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1085YA24"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end cloned, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      194 a      267 c      195 g      313 t      1 others
ORIGIN
Query Match      38.8%; Score 456; DB 13; Length 970;
Best Local Similarity 93.8%; Pred. No. 4.9e-73;
Matches 486; Conservative 0; Mismatches 30; Indels 2; Gaps 1;
QY 648 ACTGTGATTCGGTGTCCCAAGACACTGAGGTGGTATTATATATCCAGTGTATTATCA 707
Db 9 ACTGTGATTCGGTGTCCCA--GACCTGGAGGTGGTATTATATATCCAGTGTATTATCA 66
QY 708 AGTTTACCGAGTTTAGGTTTTCATTGCTACTTGGGCAAGTTGACTTGGACTAGCA 767
Db 67 AGTTTACCGAGTTTAGGTTTTCATTGCTACTTGGGCAAGTTGACTTGGACTAGCA 126
QY 768 GGCACCATGCTTCTGACCGCGAGCTGTACTTCTTACATACGCTGCTGCTGCTGCTGCTG 827
Db 127 GGCACCATGCTTCTGACCGCGAGCTGTACTTCTTACATACGCTGCTGCTGCTGCTGCTG 186
QY 828 CGTTTGTGTGGTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
Db 187 CGTTTGTGTGGTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
QY 888 TTTTCGTATTCATTTCTAAAGAAATCTGAAAGAGAGAGACAAAGAAAGAACTGAGACA 947
Db 247 TTTTCGTATTCATTTCTAAAGAAATCTGAAAGAGAGAGACAAAGAAAGAACTGAGACA 306
QY 948 GAAAGTGGAAATGAAACTCCGCTACATTCAGATGAACAAAGAACCCACAGACACCGCT 1007
Db 307 GAAAGTGGAAATGAAACTCCGCTACATTCAGATGAACAAAGAACCCACAGACACCGCT 366
QY 1008 TCTCTCCCTCCCAATCCTGCTGATCCAGTATCCCTGACAAAGAAACAGTACTGTGGC 1067
Db 367 TGTCTGCTTCCCAATCCTGCTGATTCAGTATCCCTGACAAAGAAACAGTACTGTGGG 426
QY 1068 CCTCTCACCAGCGGGTGTATCAACGCTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCT 1127
Db 427 CCTCTCACCAGTGGTTCTCTCAACGCTTCAACCCAGGCGCAGTAACTCAAGTCACTTCA 486
QY 1128 TTTAATCTGGCCAGTCTCTGAGAGAGTTCAGTAAATACAAAC 1165
Db 487 TTTAATCTGGCCAGTCTCTGAGAGAGTTCAGTAAATACAAAC 524

RESULT 6
BX370209 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
LOCUS clone CS0D1085YA24 5-PRIME, mRNA sequence.
DEFINITION
ACCESSION BX370209.1 GI:26103922
VERSION BX370209
KEYWORDS HVC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
COMMENT 1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

```

99279253  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, Y., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (KISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 50,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1510)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuritara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Sakai, K., Sakazume, N., Sogabe, Y., Tagami, M., Tanaka, K., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222).

Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

FEATURES  
source  
1. .1510  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:E030016M12"  
/db\_xref="taxon:10090"  
/clone="E030016M12"  
/tissue="lung"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
misc\_feature  
1. .1510  
/note="hypothetical Immunoglobulin structure containing protein (SCOP|48726, evidence: SCOP)"  
BASE COUNT 401 a 364 c 384 g 361 t  
ORIGIN  
Query Match 28.0%; Score 328.6; DB 11; Length 1510;  
Best Local Similarity 61.2%; Pred. No. 6.5e-50;  
Matches 636; Conservative 0; Mismatches 339; Indels 64; Gaps 4;  
QY 24 GAGATGGGGTCTGTGATTTTCCTCCAGGTTCTGGGCTGGTAAATGAAGTCATAGAGGC 83  
Db 136 GTGCTGTGTCATCTCTGGCTGAGCTGACAGCTTCGGATCCAGTTATCAGATCATAGAGGT 195  
QY 84 CCCAGAAATCAACAGTCTCTGAAGGGTCCAGGCTCGCTTCAACTGCACCGTCTCCAG 143  
Db 196 CCTCAGAAATCAACAGTCTCTGAAGGGTCCAGGCTCGCTTCAACTGCACCGTCTCCAG 255  
QY 144 GGCTGGAGCTCATCATCTGGGCTCTCAGTGACATGCTGCTGCTGCTGCTGCTGCTGCTG 203  
Db 256 GGCTGGAGCTCTCTCATCTGGGCTCTTAAACAAATGGTGGTGGTGGTGGTGGTGGTGG 315  
QY 204 GAGCCATCATCAATCAATGACCGCTTCACTCTCAGAGGTACCAAGGGGGGGAACCTTC 263  
Db 316 GGACCATCATCAACCAACCAACCGCTTCACTCTGCGAGTTACACAGCACTGACAGTTC 375  
QY 264 ACCTCGAGATGATCATCAATGAGCCAGCTGAGTTCGGGGGAACATCAGATGAGC 323  
Db 376 ATCTCGGAGTTGATCATCCATGATGTGACGCCAGTACCTCGGGATCCGTGCAATGAGC 435  
QY 324 CTCAGAAACAGTCCCTGCGATGATCTGCTTACCTTACCTTCAAGTTATGGGAGAGCTG 383  
Db 436 CTGAGAAACAGCAATGGGTTGGATCTGCTTCTCTCAGTGAGTCATGGGAGAGCCTG 495  
QY 384 TTCAATCCAGTGTAACTTTGTAGTCGCTGAGAAATGAACCTTTGTGAAGTACTTGTCTA 443  
Db 496 AACATTCCTAGCAACACCTTATAGTCACTGAGGGTGAACCCCTGTAATGTGACTTGTCT 555  
QY 444 CCCTCACTGAGCCCGGCTCCCGGATATTTCTTGGAGGCTGGTCTCTCTGCTGAGCCAT 503  
Db 556 GCCTGGGCTGAGCCCTCACTCCCGGATATTTCTTGGAGGCTTGGGTTCCCGTAAGCAT 615  
QY 504 TCAAGCTATTATTATTGTTGTTCCGGAGCC-CAGCGACCTTCAAGTGCAGTGAGCACTCTGGC 562  
Db 616 TCGAGTTACATTCCTTCTTGGAGCCGGGCACTTTATGAGGGTCTTGTAGTCTCTGGA 675  
QY 563 TCTGACCCCAAGAGCAATGGGCTTTGACTTGGTGGCTTACCTGGAGAGGCTGAGAGGC 622  
Db 676 CCTCACCACTGGGCAACCGGACCTTGACTTGTGGCAGAGCTGAAGCACTTGCAGGC 735  
QY 623 CCGCAAGTCTGCACTCTGAATCTCACTGTGATTCGGTGTCCCAAGACACTGAGGTGG 682  
Db 736 CAGCAAGTCTCTTAACTGTCAACCTGCTGTTGGTTCAGCTCCACCTGACAGTATTGGAGA 795



cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM9844 row: n column: 10  
 High quality sequence stop: 694.

## FEATURES

Location/Qualifiers  
 1. 794  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4236489"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI CGAP Kid14"  
 /note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.75 kb. Constructed by Life  
 Technologies. Note: this is a NCI CGAP Library. |"  
 BASE COUNT 186 a 214 c 215 g 179 t  
 ORIGIN

Query Match 23.7%; Score 279; DB 10; Length 794;  
 Best Local Similarity 68.3%; Pred. No. 7.2e-41;  
 Matches 416; Conservative 0; Mismatches 190; Indels 3; Gaps 2;  
 24 GAGATGGGGCTGTGATTTTCTCCACGGTTCTGGGCTCGGTAAATGAGTCATAGAGGC 83  
 123 GTCTGTGTCATCTCTGGCTCAGCTGACAGCTTCGGATCCAGTTATCAGATCATAGAGGT 182  
 84 CCCAGAAATGCAAGCTCTGAAGGGCTCCAGGCTCGCTTCAACTGCACCGTCTCCAG 143  
 183 CCTCAGAATGTAACAGCTCTAAAGGACTCAGAGGCTCACTTCAACTGCACCGTGAATCAC 242  
 144 GGCTGGAAGTCTCATGTGGGCTCTCAGTGACATGGTGTGTCTAAGCGTCAGGCCATG 203  
 243 GGCTGGAGGCTTCTCATGTGACATCTTAACCAATGGTGTGTCTGAGTCTCACCACCAA 302  
 204 GAGCCCATCATCAACAAATGACCGCTTCACTCTCAGAGGTACGACCGGGCGGAACTTC 263  
 303 GGACCCATCATCAACCAACCGCTTCACTATGCGAGTTACAACAGCACTGACAGCTTC 362  
 264 ACTCGAGATGATCATCAATGAGGAGCGCCAGTATCGGGGAACTCAGATGACGAC 323  
 363 ATCTCGAGTGTATCATCCATGATGTCAGCCGAGTGCCTGGATCCGTGCAATGACG 422  
 324 CTCAGAACAGTCCGCTGCAATGATCTGTACCTTACCGTCCAAAGTTATGGGAGAGCTG 383  
 423 CTGCAGAACAGCCATGGGTTGGATCTGCCTTCTCTCAGTCAAGTCAATGGGAGCCCTG 482  
 384 TTCAATCCCAAGTGTAACTTGTAGTGTGAGTGAATGACCTGTGAACTTACTTCTTA 443  
 483 AACATCTTAGCAACACCTTATAGTACATGAGGGTGAACCTGTATGATGATCTCTAT 542  
 444 CCCTCACTGAGACCGGCTCCCGATATTTCTGGGAGCTCGGTCTCTCTGGTTCAGCCAT 503  
 543 GCGTGGGCTGGACCTCACTCCCGATATTTCTGGGAGCTTGAGGTTCCCGTAAGCCAT 602  
 504 TCAAGCTATTATTTGTTCCGGAGCCAGGACCTTCAAGTGCAGTGAATCTCTGGCT 563  
 603 TCGAGTTTACAATTCTTTCTGGAGCC--GGGCACTTTATGAGGCTCTGAGTGTCTCTGAA 660  
 564 CTGACCCCAACAGAGCAATGGGCTTTGACCTTGGTGGCTACCTCGGAGAGCCCTGAGGCC 623  
 661 CTCA-CACACTGGCAAGCGGACCTTGACTTGTGTGCAAGAGCTGAGAGCTTGCAGGC 719  
 624 CGCAAGTCT 632  
 720 CAGCAGTCT 728

## RESULT 9

BC740428  
 LOCUS 725 bp mRNA linear EST 15-MAY-2001  
 DEFINITION 602633817F1 NCI\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4778789 5';  
 mRNA sequence.  
 ACCESSION BC740428  
 VERSION BC740428  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 725)  
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: James Cleaver, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA  
 Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10635 row: b column: 06  
 High quality sequence stop: 725.

## FEATURES

Location/Qualifiers  
 1. 725  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="IMAGE:4778789"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI CGAP Skn3"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. primer: Oligo dt.  
 Average insert size 1.5kb. Library constructed by Life  
 Technologies. Note: this is a NCI CGAP Library. |"  
 BASE COUNT 227 a 163 c 159 g 176 t  
 ORIGIN

Query Match 22.8%; Score 267.4; DB 10; Length 725;  
 Best Local Similarity 99.6%; Pred. No. 9.3e-39;  
 Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 906 AAGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAAC 965  
 Db AGGAAATCTGAAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAAC 436  
 QY 966 TCCGCTACAAATTCAGATGAACAAAGACACACACCGCTTCTCTCCCTCCCAATCC 1025  
 Db 437 TCCGCTACAAATTCAGATGAACAAAGACACACACCGCTTCTCTCCCTCCCAATCC 496  
 QY 1026 TGTGAATCCAGTATCTCTGAAACAAAGAAAGTGTGGCCCTCTCTCACCAGCGGCT 1085  
 Db 497 TGTGAATCCAGTATCTCTGAAACAAAGAAAGTGTGGCCCTCTCTCACCAGCGGCT 556  
 QY 1086 GATCAAGTCCACCGGCGAGAGTCAATCCACAGGCTCTTTTAACTCTGGCAGTCTCT 1145  
 Db 557 GATCAAGTCCACCGGCGAGAGTCAATCCACAGGCTCTTTTAACTCTGGCAGTCTCT 616  
 QY 1146 GAGAAGTCAAGTAACTAATCTGTAGTATA 1174  
 Db 617 GAGAAGTCAAGTAACTAATCTGTAGTATA 645

## RESULT 10

BC370210  
 LOCUS 1148 bp mRNA linear EST 08-MAY-2003  
 DEFINITION BX370210 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS01085YA24 5-PRIME, mRNA sequence.  
 ACCESSION BX370210

```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

BX370210.1 GI:30451881  
 EST.  
 Homo sapiens (human)  
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 1148)  
 Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 10299.f, and  
 it belongs to a clone representative of this cluster. For more  
 information about this cluster and the virtual cDNA, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0BAG039ZE05\_CS03685\_2&cluster=10299.f.  
 Contact : Rong Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0BAG039ZE05\_CS03685\_2.  
 Location/Qualifiers  
 1..1148  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1085YA24"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 204 a 263 c 137 g 544 t  
 ORIGIN

Query Match 22.7%; Score 266.4; DB 13; Length 1148;  
 Best Local Similarity 89.2%; Pred. No. 1.3e-38;  
 Matches 321; Conservative 0; Mismatches 36; Indels 3; Gaps 3;

```

QY 651 GTGATTGGTGTCCCAAGACACTGGAGTGGTATTAATATCCAGTGTATATCAAGT 710
    |||||
Db 1 GTGATTGGTGT-TCCAAGACACTGGAGTGGTATTAATATCCAGTGTATATCAAGT 59
    |||||

QY 711 TTACCGAGTTAGTGTTCCTTCAATTCCTTCTTGGGCAAAAGTTGGACTTGGACTAGCAGC 770
    |||||
Db 60 TTACCGAGTTAGTGTTCCTTCAATTCCTTCTTGGGCAAAAGTTGGACTTGGACTAGCAGC 119
    |||||

QY 771 ACCATGCTTCTGACGCGCGTCTACTCTTACAAATAGCTGCTGCTGCTGCGCGCGTCT 830
    |||||
Db 120 ACCATGCTTCTGACGCGCGTCTACTCTTACAAATAGCTGCTGCTGCTGCGCGCGTCT 179
    |||||

QY 831 TGTGTGGTGCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
    |||||
Db 180 TGTGTGGTGCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
    |||||

QY 891 CGTATTCATTTTCAAGAAAT-TCTGAAAGAGAGACACAAACAAAGAACTGACACAG 949
    |||||
Db 240 TGTATTTATATTTAAAGAAATTTTGTAAAGTGAAGACACTCAGTGAACAACTGTGACAG 299
    |||||

QY 950 AAG-TGGAATGAAATCCCGGTACAAATTCAGATGAAACAAAGACACACACCGCTT 1008
    |||||
Db 300 AAGTTGTTAATGAAATCTCGTTTTCATTTTCATTTGTTCTTTTGACACAGTAATCCGTT 359
    |||||

```

RESULT 11  
 BB663870  
 LOCUS  
 DEFINITION BB663870 RIKEN full-length enriched, 0 day neonate lung Mus  
 musculus cDNA clone E030016M12 5', mRNA sequence.  
 BB663870  
 ACCESSION

```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

BB663870.1 GI:16497624  
 EST.  
 Mus musculus (house mouse)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 631)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,  
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ono, M., Sasaki,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,  
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
 Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa,  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
 Location/Qualifiers  
 1..631  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="E030016M12"  
 /tissue\_type="lung"  
 /dev\_stage="0 day neonate"  
 /lab\_host="DHL0B"  
 /clone\_lib="RIKEN full-length enriched, 0 day neonate  
 lung"  
 /note="Site 1: Sall; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGCGCCGCAACTCGAGTTTTTTTTTTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5'

USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68913-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: [smith@email.marc.usda.gov](mailto:smith@email.marc.usda.gov)

## CONCLUSION



QY 324 CTCAGAACAGTCGCCTGCATGATCTGCTTACCTTACCGTCCAGTTATGGAGAGCTG 383  
 Db 463 CTCAGAACAGTCGCCTGCATGATCTGCTTACCTTACCGTCCAGTTATGGAGAGCTG 522  
 QY 384 TTCATTCCTCCAGTGTGTTAATCTTGTAGTGGCTGAG 416  
 Db 523 AACATTCCTAGCAACACCTTATAGTCACITGAG 555

RESULT 15  
 BE476432 517 bp mRNA linear EST 27-MAR-2003  
 LOCUS 159670 BARS 5BOV Bos taurus cDNA 5', mRNA sequence.  
 DEFINITION BE476432  
 ACCESSION BE476432  
 VERSION BE476432.1 GI:9595965  
 KEYWORDS EST.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus

REFERENCE AUTHORS  
 Sonstegard, T., Capuco, A.V., White, J., Van Tassel, C.P., Connor, E.E.,  
 Cho, J., Sultana, R., Shade, L., Wray, J.E., Wells, K.D. and  
 Quackenbush, J.  
 TITLE Analysis of bovine mammary gland EST and functional annotation of  
 the Bos taurus gene index  
 JOURNAL Mamm. Genome 13 (7), 373-379 (2002)  
 MEDLINE 22135956  
 PUBMED 12140684

COMMENT  
 Contact: Sonstegard TS  
 USDA, ARS, Beltsville Agricultural Research Center  
 Bldg. 200 Rm 2A, Beltsville, MD 20705, USA  
 Tel: 301 504 8416  
 Fax: 301 504 8414  
 Email: tads@psi.barc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCACGACG  
 Plate: 27 row: H column: 19  
 Seq primer: ATTTAGGTGACACTATAG.

FEATURES  
 source  
 1..517  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /clone\_lib="BARS 5BOV"  
 /note="Vector: pCMV SPORTE6; Site 1: NotI; Site 2: SalI;  
 Library made from pooled mRNA isolated from mammary  
 tissues at eight physiological, developmental, and disease  
 states."  
 BASE COUNT 117 a 133 c 165 g 102 t  
 ORIGIN

Query Match 16.9%; Score 199; DB 10; Length 517;  
 Best Local Similarity 73.8%; Pred.No. 2.6e-26;  
 Matches 253; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 24 GAGATGGGGCTGTGATTTTCTCCAGGTTCTGGGTCGTGTAATGAAGTCATAGAAGGC 83  
 Db 175 GTGCTGGTGGCTCTGGCGCTGGCAGCTGTGGATCCAGCAGTGAATCATAGAGGT 234  
 QY 84 CCCAGATGAACAGTCCTGAGGGCTCCAGGCTCGCTTCACTGCACCGTCTCCAG 143  
 Db 235 CCCAAGAATGTACAGAGGCTGAGGGCTCGGCTTCACTGCACCATCTCCAG 294  
 QY 144 GGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGTGTGTTAAGCGTCAGGCCCATG 203

Db 295 GGCTGGAAGCTCTGTCATGTGGGCTCTGAGAGGCACAGTGGTGTGATGACACCTAAT 354  
 QY 204 GAGCCCATCATCACCACCAATGACCGCTTCACTCTCAGAGGTACGACCGGGGGAACCTTC 263  
 Db 355 GAGACCATCATCACCACCAATGACCGCTTCACTCTCAGAGGTACGACCGGGGGAACCTTC 414  
 QY 264 ACCTCGGAGATGATCATCCCAATGTGGAGCCCAAGTGTTCGGGGAACATCAGATGCAGC 323  
 Db 415 ATCTCTGAGATGATAATTTCATGACGTGCAACTGAGCGATGCCGGCAAGTCAATGCAGC 474  
 QY 324 CTCAGAACAGTCGCCTGCATGATCTGCTTACCTTACCGTCC 366  
 Db 475 CTTCAGAACAGTCGCCTGCATGATCTGCTTACCTTACCGTCC 517

Search completed: November 12, 2003, 23:19:35  
 Job time : 2822.33 secs





## US-09-491-356C-14

Query Match 3.8%; Score 44.8; DB 4; Length 265;  
Best Local Similarity 81.2%; Pred. No. 0.0031;  
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 808 GCTGCTGCTGCTGCGCGCGTGTGTGTGGCTGCAACTGCTGCTGCGGTGTGTCT 867  
DB 158 GCTGTGCTGCTGCTGCTGCTGTGTGTGTGCTGCTGCTGCTGCTGCTGCT 99  
QY 868 GCTG 871  
DB 98 GCTG 95

## RESULT 6

US-09-491-356C-16/c  
; Sequence 16, Application US/09491356C  
; Patent No. 6566061  
; GENERAL INFORMATION:  
; APPLICANT: Philibert, Robert A.  
; APPLICANT: Ginns, Edward I.  
; APPLICANT: Delisi, Lynn  
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13  
; FILE REFERENCE: 9465.6US11  
; CURRENT APPLICATION NUMBER: US/09/491,356C  
; CURRENT FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: PCT/US99/09365  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 60/083,465  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 265  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-491-356C-15

Query Match 3.8%; Score 44.8; DB 4; Length 265;  
Best Local Similarity 81.2%; Pred. No. 0.0031;  
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 808 GCTGCTGCTGCTGCGCGCGTGTGTGTGGCTGCAACTGCTGCTGCGGTGTGTCT 867  
DB 158 GCTGTGCTGCTGCTGCTGCTGTGTGTGTGCTGCTGCTGCTGCTGCTGCT 99  
QY 868 GCTG 871  
DB 98 GCTG 95

## RESULT 7

US-09-491-356C-17/c  
; Sequence 17, Application US/09491356C  
; Patent No. 6566061  
; GENERAL INFORMATION:  
; APPLICANT: Philibert, Robert A.  
; APPLICANT: Ginns, Edward I.  
; APPLICANT: Delisi, Lynn  
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13  
; FILE REFERENCE: 9465.6US11  
; CURRENT APPLICATION NUMBER: US/09/491,356C  
; CURRENT FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: PCT/US99/09365  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 60/083,465  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 265  
; TYPE: DNA

; ORGANISM: Homo sapiens  
US-09-491-356C-17

Query Match 3.8%; Score 44.8; DB 4; Length 265;  
Best Local Similarity 81.2%; Pred. No. 0.0031;  
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 808 GCTGCTGCTGCTGCGCGCGTGTGTGTGGCTGCAACTGCTGCTGCGGTGTGTCT 867  
DB 158 GCTGTGCTGCTGCTGCTGCTGTGTGTGTGCTGCTGCTGCTGCTGCTGCT 99  
QY 868 GCTG 871  
DB 98 GCTG 95

## RESULT 8

US-09-491-356C-18/c  
; Sequence 18, Application US/09491356C  
; Patent No. 6566061  
; GENERAL INFORMATION:  
; APPLICANT: Philibert, Robert A.  
; APPLICANT: Ginns, Edward I.  
; APPLICANT: Delisi, Lynn  
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13  
; FILE REFERENCE: 9465.6US11  
; CURRENT APPLICATION NUMBER: US/09/491,356C  
; CURRENT FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: PCT/US99/09365  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 60/083,465  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 265  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-491-356C-18

Query Match 3.8%; Score 44.8; DB 4; Length 265;  
Best Local Similarity 81.2%; Pred. No. 0.0031;  
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 808 GCTGCTGCTGCTGCGCGCGTGTGTGTGGCTGCAACTGCTGCTGCGGTGTGTCT 867  
DB 158 GCTGTGCTGCTGCTGCTGCTGTGTGTGTGCTGCTGCTGCTGCTGCTGCT 99  
QY 868 GCTG 871  
DB 98 GCTG 95

## RESULT 9

US-09-491-356C-2/c  
; Sequence 2, Application US/09491356C  
; Patent No. 6566061  
; GENERAL INFORMATION:  
; APPLICANT: Philibert, Robert A.  
; APPLICANT: Ginns, Edward I.  
; APPLICANT: Delisi, Lynn  
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13  
; FILE REFERENCE: 9465.6US11  
; CURRENT APPLICATION NUMBER: US/09/491,356C  
; CURRENT FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: PCT/US99/09365  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 60/083,465  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 6794

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-491-356C-2

Query Match      3.8%; Score 44.8; DB 4; Length 6794;
Best Local Similarity 76.4%; Pred. No. 0.017;
Matches 55; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCCCGCGTGTGTGGTGGCTGCAACTGCTGCGCCGTTGTTTCT 867
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6349 GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867

QY 868 GCTGTAGAGAA 879
      ||||| ||||| |||||
Db 6289 GCTGCTGCCGA 6278

```

```

RESULT 10
US-09-491-356C-1/c
; Sequence 1, Application US/09491356C
; Patent No. 656061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Ginns, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491,356C
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 55298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (485)..(485)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (838)..(838)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (16728)..(16728)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22750)..(22750)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22756)..(22756)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (28519)..(28519)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (44804)..(44804)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (45002)..(45002)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54049)..(54049)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54226)..(54226)
; OTHER INFORMATION: n is not determined
US-09-491-356C-1

```

Query Match 3.8%; Score 44.8; DB 4; Length 55298;

```

Best Local Similarity 81.2%; Pred. No. 0.05;
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCCCGCGTGTGTGGTGGCTGCAACTGCTGCGCCGTTGTTTCT 867
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23944 GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 23885

QY 868 GCTG 871
      |||||
Db 23884 GCTG 23881

RESULT 11
US-09-620-312D-475/c
; Sequence 475, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL genes Version 1.0
; SEQ ID NO 475
; LENGTH: 3302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(2563)
US-09-620-312D-475

```

```

Query Match      3.7%; Score 43.4; DB 4; Length 3302;
Best Local Similarity 69.4%; Pred. No. 0.029;
Matches 59; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCCCGCGTGTGTGGTGGCTGCAACTGCTGCGCCGTTGTTTCT 867
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 527 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468

QY 868 GCTGTAGAGAAAAGAGGATTCG 892
      ||||| ||||| |||||
Db 467 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 443

```

```

RESULT 12
US-09-086-663A-70/c
; Sequence 70, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA

```

	Query Match Best Local Similarity Matches 51; Conservative	3.7%; Score 43.2; DB 4; Length 3334; Pred. No. 0.033; Mismatches 0; Indels 0; Gaps 0;
Qy	808 GTCGCTGCTGCCGCGTGTGTGTGCTGCTGCAACTGCTGCTGCCCTTCTTCTTCT 867	
Db	534 GTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475	
Qy	868 GCTG 871	
Db	474 GCTG 471	

RESULT 15  
US-08-531-927B-1/c  
Sequence 1, Application US/08531927B  
Patent No. 5840491  
GENERAL INFORMATION:  
APPLICANT: Kakizuka, Akira  
TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph  
Patent No. 5840491  
TITLE OF INVENTION: Disease Gene and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173-4799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,927B  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP H6-251600  
FILING DATE: 21-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: ATH95-01A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240

```

; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1776 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1115
US-08-531-927B-1

Query Match      3.6%; Score 42.8; DB 2; Length 1776;
Best Local Similarity 61.8%; Pred. No. 0.031;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCCCGCGTGTGTTGGCTGCACACTGCTGCTGCCGTTGTTCT 867
Db 975 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
QY 868 GCTGTAGAGAGAGAGAGGATTTCGTATTCAATTCAGAGAGAAATCTGAA 917
Db 915 GCTGCTGTTTTTCAAGTAGGCTTCCTGCTCTTCGAGAGCTCTTCTGAA 866

```

Search completed: November 12, 2003, 23:23:36  
Job time : 79.7212 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2003, 20:56:43 ; Search time 393.853 Seconds  
(without alignments)  
9519.735 Million cell updates/sec

Title: US-09-729-264-1  
Perfect score: 1175  
Sequence: 1 cgtgtgcccactgaataa.....gtaatacaactgtagttag 1175

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2141354 seqs, 1595478879 residues  
Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/prodata/1/pubna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/prodata/1/pubna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/prodata/1/pubna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/prodata/1/pubna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/prodata/1/pubna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/prodata/1/pubna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/prodata/1/pubna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/prodata/1/pubna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/prodata/1/pubna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/prodata/1/pubna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/prodata/1/pubna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/prodata/1/pubna/US09C\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/prodata/1/pubna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/prodata/1/pubna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/prodata/1/pubna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/prodata/1/pubna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/prodata/1/pubna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	269	22.9	474	11	US-09-918-995-3342
2	190.4	16.2	401	9	US-09-864-761-16305
3	142	12.1	398	10	US-09-983-965-4945
4	46	3.9	3577	13	US-10-008-739A-1
5	44.8	3.8	1310	10	US-09-849-243-13
6	44.8	3.8	2614	11	US-09-822-846-491
7	44.8	3.8	3263	10	US-09-849-243-15
8	44.8	3.8	4286	10	US-09-849-243-14
9	44.8	3.8	5085	14	US-10-198-846-9854
10	44.8	3.8	6604	10	US-09-880-107-1748
11	44.2	3.8	2790	12	US-10-029-386-2626
12	44	3.7	299	12	US-09-814-353-4844
13	44	3.7	299	12	US-09-814-353-1141
14	44	3.7	385	12	US-09-814-353-5368
15	44	3.7	385	12	US-09-814-353-11655
16	44	3.7	455	10	US-09-728-444-151

17	43.6	3.7	698	12	US-10-029-386-22908
18	43.6	3.7	4321	14	US-10-205-823-35
19	43.6	3.7	4393	14	US-10-198-846-10819
20	43.4	3.7	384	10	US-09-738-626-544
21	43.4	3.7	3302	14	US-10-037-270-475
22	43.4	3.7	3309400	10	US-09-738-626-1
23	43.4	3.7	293	9	US-09-864-761-18923
24	43.2	3.7	459	9	US-09-864-761-2182
25	43.2	3.7	1791	9	US-09-416-384A-6
26	42.8	3.6	586	12	US-10-029-386-9094
27	42.8	3.6	4248	14	US-10-198-846-11546
28	42.8	3.6	3673778	12	US-10-312-841-1
29	42.6	3.6	388	12	US-09-814-353-17525
30	42.6	3.6	797	12	US-09-814-353-4971
31	42.6	3.6	797	12	US-09-814-353-11264
32	42.6	3.6	2424	13	US-10-029-217A-25
33	42.6	3.6	4959	13	US-10-029-217A-1
34	42.6	3.6	4960	13	US-10-029-217A-30
35	42.4	3.6	1925	14	US-10-219-220-217
36	42.4	3.6	2453	12	US-10-241-009-22
37	42.4	3.6	2453	12	US-10-190-434B-22
38	42.2	3.6	5257	12	US-10-007-926A-237
39	42.2	3.6	5959	10	US-09-954-456-1396
40	42	3.6	503	12	US-09-814-353-5281
41	42	3.6	503	12	US-09-814-353-11568
42	42	3.6	539	12	US-09-814-353-17862
43	42	3.6	632	12	US-09-814-353-5078
44	42	3.6	632	12	US-09-814-353-11370
45	41.8	3.6	3673778	12	US-10-312-841-2

## ALIGNMENTS

### RESULT 1

US-09-918-995-3342  
; Sequence 3342, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3342  
; LENGTH: 474  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(474)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-3342

Query Match	22.9%	Score 269;	DB 11;	Length 474;
Best Local Similarity	98.2%;	Pred. No. 6.8e-71;		
Matches 272;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	899	ATTTCAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGACACAGAAAGTGGAAA	958	
DB	54	ACTTTTAAAGAAATCTGAAAAGAGAGACAAACAAAGAACTGACACAGAAAGTGGAAA	113	
QY	959	TGAAATCTCGGCTACAAATTCAGATGACAAAAGACACACAGACACCGCTTCTCCCTCC	1018	
DB	114	TGAAATCTCGGCTACAAATTCAGATGACAAAAGACACACAGACACCGCTTCTCCCTCC	173	
QY	1019	CAATCTCTGTGAATCCAGTGTATCTTGAACAAAGAAACAGTAGCTGTGGCCCTCTCACC	1078	

Db 174 CAATCTCTGTAATCCAGTGTCTCTGACAAAGAAACAGTAGTGTGGCCCTCTCACCA 233  
 QY 1079 GCGGCTGATCAACGCTCCACCCAGCCAGCAGCAAGTCAATCCAGGCTTCTTTTAATCTGGC 1138  
 Db 234 GCGGCTGATCAACGCTCCACCCAGCCAGCAGCAAGTCAATCCAGGCTTCTTTTAATCTGGC 293  
 QY 1139 CAGTCTGAGAGGTCACTAATACACTGTAGTATAG 1175  
 Db 294 CAGTCTGAGAGGTCACTAATACACTGTAGTATAG 330

RESULT 2  
 US-09-864-761-16305  
 ; Sequence 16305, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; FILE REFERENCE: Aemica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 16305  
 ; LENGTH: 401  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AF121782.1  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9  
 US-09-864-761-16305

Query Match 16.2%; Score 190.4; DB 9; Length 401;

Best Local Similarity 97.0%; Pred. No. 4.6e-47;  
 Matches 194; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 42 TTCTTCCACGGTTCCTGGTCTGTAATGAAGTCAATAGAGGCCCCAGAAATGCAACAGTC 101  
 Db 202 TACCTTCCAGGTTCTGGGTCTGTAATGAAGTCAATAGAGGCCCCAAATGCAAGAGTC 261  
 QY 102 CTGAAGGGCTCCAGGCTCGTTCAACTGCAAGGCTCCAGGCTGGAAGCTCATCATG 161  
 Db 262 CTGAAGGGCTCCAGGCTCGTTCAACTGCAAGGCTCCAGGCTGGAAGCTCATCATG 321  
 QY 162 TGGGCTCTCAGTGACATGCTGCTTAAGCGTTCAGGCCCCATGAGCCCCATCATCAAT 221  
 Db 322 TGGGCTCTCAGTGACATGCTGCTTAAGCGTTCAGGCCCCATGAGCCCCATCATCAAT 381  
 QY 222 GACCGCTTCACCTCTCAGAG 241  
 Db 382 GACCGCTTCACCTCTCAGAG 401

RESULT 3  
 US-09-983-965-4945  
 ; Sequence 4945, Application US/09983965  
 ; Patent No. US20020137160A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Warren, Wesley C.  
 ; APPLICANT: Tao, Nengbing  
 ; APPLICANT: Byatt, John C.  
 ; APPLICANT: Mathialagan, Nagappan  
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 ; FILE REFERENCE: 37-21(10297)C  
 ; CURRENT APPLICATION NUMBER: US/09/983,965  
 ; CURRENT FILING DATE: 2001-10-26  
 ; PRIOR APPLICATION NUMBER: US 09/465,231  
 ; PRIOR FILING DATE: 1999-12-15  
 ; PRIOR APPLICATION NUMBER: US 60/113,678  
 ; PRIOR FILING DATE: 1998-12-17  
 ; NUMBER OF SEQ ID NOS: 5912  
 ; SEQ ID NO 4945  
 ; LENGTH: 398  
 ; TYPE: DNA  
 ; ORGANISM: Bos taurus  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 26-LIB34-017-Q1-E1-G9  
 US-09-983-965-4945

Query Match 12.1%; Score 142; DB 10; Length 398;  
 Best Local Similarity 74.8%; Pred. No. 2.3e-32;  
 Matches 178; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
 QY 24 GAGATGGGGCTGTGTGATTTTCTCCACGGTTCTGGTCTGGTAATGAAGTCAATAGAGC 83  
 Db 161 GTGCTGTGTCTCTGCGCTGTGGCAGCTGTGGATCCAGCAGTGAATATCATAGAGGT 220  
 QY 84 CCCAGAAATGCAACAGTCTGGAAGGCTCCAGGCTCGCTTCACTGCAACCTCTCCAG 143  
 Db 221 CCCAAGAAATGTACAGCCCTGGAAGGCTCGGAGGCTCGCTTCACTGCAACCTCTCCAG 280  
 QY 144 GGCTGGAAGCTCATCATGTGGGCTCTCAGTGACATGGTGTGCTAAGCGTCAAGGCCCATG 203  
 Db 281 GGCTGGAAGCTCGTCTCATGTGGGCTCTGAGAGGCACAGTGTGCTGAGCATGACACCTAAT 340  
 QY 204 GAGCCCATCATCACCAATGACCGCTTCACTCTCAGAGGTAGCAGCCGCGGAACT 261  
 Db 341 GAGACCATCATCACCAAGTACCGCTTCACTCTCGGCAAGCTTACCAAGAGGCGCGAACT 398

RESULT 4  
 US-10-008-739A-1/c  
 ; Sequence 1, Application US/10008739A  
 ; Publication No. US20020161194A1  
 ; GENERAL INFORMATION:

```
; APPLICANT: Pfizer Inc.
; APPLICANT: Castleberry, Tessa A.
; APPLICANT: Lu, Bihong
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: The Canine Androgen Receptor
; FILE REFERENCE: PC10893AGR
; CURRENT APPLICATION NUMBER: US/10/008,739A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3577
; TYPE: DNA
; ORGANISM: Canine
US-10-008-739A-1

Query Match      3.9%; Score 46; DB 13; Length 3577;
Best Local Similarity 65.7%; Pred. No. 0.014;
Matches 67; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 776 GCTTCTGACGCGAGTGTAATCTTACAAATAGCTGCTGCTGCTGCCCGCGTGTGTG 835
Db 633 GCTGCTGCTACCTTCTGATACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 574
QY 836 TGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877
Db 573 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 532

RESULT 5
US-09-849-243-13/c
; Sequence 13, Application US/09849243
; Patent No. US2002015127A1
; GENERAL INFORMATION:
; APPLICANT: Kirschbaum, Bernd
; Meisler, Erik
; Polites, Greg
; TITLE OF INVENTION: PURIFICATION OF HIGHER ORDER TRANSCRIPTION
; COMPLEXES FROM TRANSGENIC
; NON-HUMAN ANIMALS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSER: HELLER, EHRMAN, WHITE & McAULIFFE
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/849,243
; FILING DATE: 07-May-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Granados, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 38005-0148
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)912-2000
; TELEFAX: (202)912-2020
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

; APPLICANT: Pfizer Inc.
; APPLICANT: Castleberry, Tessa A.
; APPLICANT: Lu, Bihong
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: The Canine Androgen Receptor
; FILE REFERENCE: PC10893AGR
; CURRENT APPLICATION NUMBER: US/10/008,739A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3577
; TYPE: DNA
; ORGANISM: Canine
US-10-008-739A-1

Query Match      3.9%; Score 46; DB 13; Length 3577;
Best Local Similarity 65.7%; Pred. No. 0.014;
Matches 67; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 776 GCTTCTGACGCGAGTGTAATCTTACAAATAGCTGCTGCTGCTGCCCGCGTGTGTG 835
Db 633 GCTGCTGCTACCTTCTGATACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 574
QY 836 TGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877
Db 573 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 532

RESULT 6
US-09-822-846-491/c
; Sequence 491, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steinger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fehrel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakari
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 491
; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-846-491

Query Match      3.8%; Score 44.8; DB 11; Length 2614;
Best Local Similarity 76.4%; Pred. No. 0.027;
Matches 55; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 808 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
Db 2179 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2120
QY 868 GCTGTAGAGAA 879
Db 2119 GCTGCTGCCGA 2108

RESULT 7
US-09-849-243-15/c
; Sequence 15, Application US/09849243
; Patent No. US2002015127A1
```



Query Match	3.7%	Score 44;	DB 12;	Length 299;
Best Local Similarity	55.9%;	Mismat.No. 0.01l;		
Matches 80;	Conservative 0;	Pred.Matches 63;	Indels 0;	Gaps 0;
Qy	860	TTGTTTTTCGTCTAGAGAAAAAGAGGATTTCGTATTCAAATTCAAAAGAAATCTCGAAA	919	
Db	189	TTTTTTTTTTTTTTTTTNAAAAAAAAAAGGGNTTTTTTTTTTTTTTTTTTTTNNAAAAAAAAAAAAAAAAA	130	

```
QY 920 AGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAACTCCGGCTACAATTC 979
D 129 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 70
QY 980 AGATCAACAAAGACCAAGACAGACA 1002
D 69 AAAAAAAAAAAAAAAAAAGAAAAAAAAA 47

RESULT 13
US-09-814-353-11141/c
; Sequence 11141, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11141
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 146, 163, 175, 190, 191, 192, 195, 198, 199, 201, 202, 203,
; LOCATION: 206, 213, 214, 217, 222, 224, 225, 226, 229, 233, 239,
; LOCATION: 246, 247, 249, 254, 255, 256, 264, 268, 274, 284, 286
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-11141

Query Match 3.7%; Score 44; DB 12; Length 299;
Best Local Similarity 55.9%; Pred. No. 0.011; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 63;

QY 860 TTGTTTCTGCTAGAGAAAGAGGATTTGTTTCAATTCAGAAAGAAATCTGAAA 919
D 189 TTTTITTTTTTTTNAAGAAAGGNTTTTTTTTTTTTTTTTTTTTNAAGAAAGAA 130
QY 920 AGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAACTCCGGCTACAATTC 979
D 129 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 70
QY 980 AGATCAACAAAGACCAAGACAGACA 1002
D 69 AAAAAAAAAAAAAAAAAAGAAAAAAAAA 47

RESULT 14
US-09-814-353-5368/c
; Sequence 5368, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11141
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 146, 163, 175, 190, 191, 192, 195, 198, 199, 201, 202, 203,
; LOCATION: 206, 213, 214, 217, 222, 224, 225, 226, 229, 233, 239,
; LOCATION: 246, 247, 249, 254, 255, 256, 264, 268, 274, 284, 286
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-11141

Query Match 3.7%; Score 44; DB 12; Length 299;
Best Local Similarity 55.9%; Pred. No. 0.011; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 857 TTGTTTCTGCTAGAGAAAGAGGATTTGTTTCAATTCAGAAAGAAATCTGA 916
D 189 TTTTITGTTTNCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 130
QY 917 AAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAACTCCGGCTACA 976
D 129 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 70
QY 977 TTGATGAGACAAAGACCAAGACAGACA 1002
D 69 AAAAAAAAAAAAAAAAAAGAAAAAAAAA 44

RESULT 15
US-09-814-353-11655/c
; Sequence 11655, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5368
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 148, 149, 160, 161, 162, 163, 173, 174, 175, 176, 177, 178,
; LOCATION: 179, 181, 192, 193, 210, 213, 214, 215, 216, 218, 219, 234,
; LOCATION: 238, 239, 240, 241, 242, 245, 247, 254, 256, 261, 282, 264,
; LOCATION: 265, 267, 268, 272, 273, 274, 275, 276, 277, 278, 279
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 280, 281, 282, 283, 284, 285, 288, 289, 301, 303, 304, 305,
; LOCATION: 306, 307, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,
; LOCATION: 319, 320, 321, 322, 326, 328, 329, 341, 354, 357, 358, 361,
; LOCATION: 362, 363, 367, 369, 371, 372, 373, 374, 375, 377, 382
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 384
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-5368

Query Match 3.7%; Score 44; DB 12; Length 385;
Best Local Similarity 52.7%; Pred. No. 0.014;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
```

```
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11655
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 148, 149, 160, 161, 162, 163, 173, 174, 175, 176, 177, 178,
; LOCATION: 179, 181, 192, 193, 210, 213, 214, 215, 216, 218, 219, 234,
; LOCATION: 238, 239, 240, 241, 242, 245, 247, 254, 256, 261, 262, 264,
; LOCATION: 265, 267, 268, 272, 273, 274, 275, 276, 277, 278, 279
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 280, 281, 282, 283, 284, 285, 288, 289, 301, 303, 304, 305,
; LOCATION: 306, 307, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,
; LOCATION: 319, 320, 321, 322, 326, 328, 329, 341, 354, 357, 358, 361,
; LOCATION: 362, 363, 367, 369, 371, 372, 373, 374, 375, 377, 382
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 384
; OTHER INFORMATION: n = A,T,C or G
;
US-09-814-353-11655

Query Match      3.7%; Score 44; DB 12; Length 385;
Best Local Similarity 52.7%; Pred. No. 0.014;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy      857 TTGTTCTTTCTGCTAGAGAAAACAGAGATTTCGTTATTCATTTCAAAAGAAATCTGA 916
Db      189 TTTTGTTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 130

Qy      917 AAAAGAGACACAAACAAAGAACTGAGACAGAAAGTGGAATGAAACTCCGGCTACAA 976
Db      129 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 70

Qy      977 TTCAGATGAACAAAGACCACAGACA 1002
Db      69 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 44
```

Search completed: November 13, 2003, 03:28:10  
Job time : 400.853 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 03:28:26 ; Search time 15.227 Seconds  
(without alignments)  
2412.580 Million cell updates/sec

Title: US-09-729-264-2  
Perfect score: 2059  
Sequence: 1 MELVIFLHSGSGNEVIEGP.....HPQAFNLASPEKVSNTTVV 382

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.5	7.2	404	1 I61596	advanced glycosyla
2	133	6.5	1427	2 I51569	tumor suppressor -
3	129.5	6.3	184	2 T24272	hypothetical prote
4	129.5	6.3	344	2 I56551	neurotrophin - rat
5	128	6.2	1447	2 A54100	tumor suppressor p
6	127.5	6.2	871	1 I48696	protein-tyrosine k
7	127.5	6.2	881	1 I48696	protein-tyrosine k
8	125.5	6.1	6642	2 T29757	protein UNC-99 - C
9	124	6.0	1070	2 JC4593	protein-tyrosine k
10	122.5	5.9	108	2 T15651	hypothetical prote
11	120	5.8	345	2 JC4025	opioid-binding cel
12	119	5.8	345	2 S03199	opioid-binding pro
13	118	5.7	345	2 JC1239	opioid-binding pro
14	117.5	5.7	1443	2 I50600	neogenin - chicken
15	116.5	5.7	338	2 JC1238	opioid-binding pro
16	116	5.6	4391	2 A38096	perlecan precursor
17	114.5	5.5	152	2 T18975	hypothetical prote
18	113.5	5.5	416	1 A42879	advanced glycosyla
19	112	5.4	764	2 A49448	irregular chiasm C
20	111	5.4	693	2 S49228	sodium-dependent p
21	110.5	5.4	620	2 JH0593	Schwann cell myeli
22	110.5	5.4	1177	2 T16594	hypothetical prote
23	109.5	5.3	2491	1 A28372	insulin-like growt
24	109	5.3	391	2 T09058	butyrophilin homol
25	108.5	5.3	487	2 S65333	butyrophilin - mou
26	107	5.2	882	2 I38912	receptor tyrosine
27	107	5.2	3707	2 S18252	heparan sulfate pr
28	106.5	5.2	662	2 T16525	hypothetical prote
29	106.5	5.2	802	2 T13149	mitogen-and stress

30	106.5	5.2	946	1 A47299	ror-related recept
31	106	5.1	1092	1 JN0635	neural cell adhesi
32	104.5	5.1	862	2 I49583	differentiation an
33	104	5.1	2222	2 T13924	sdh protein - frui
34	104	5.1	5825	2 T12117	polyprotein - fava
35	103.5	5.0	890	1 A53743	protein-tyrosine k
36	102.5	5.0	333	2 A31923	amalgam protein pr
37	102.5	5.0	423	2 T29549	hypothetical prote
38	102.5	5.0	789	2 T28714	hypothetical prote
39	102.5	5.0	1273	2 T42405	sax-3 protein - Ca
40	102.5	5.0	1355	2 T28715	hypothetical prote
41	102.5	5.0	2499	1 A30788	mannose 6-phosphat
42	102	5.0	365	2 JC7780	coxackie- and ade
43	102	5.0	721	2 T41530	hypothetical prote
44	102	5.0	868	2 A46512	CD22 homolog/B lym
45	102	5.0	880	1 JC4166	protein-tyrosine k

ALIGNMENTS

RESULT 1  
I61596  
Advanced glycosylation end-products, receptor precursor - human  
N:Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprote  
C:Species: Homo sapiens (man)  
C:Date: 24-May-1996 #sequence revision 07-Feb-1997 #text\_change 16-Jul-1999  
C:Accession: I61596; B42879; S27968  
R:Shigaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko, H  
Genomics 23, 408-419, 1994  
A:Title: Three genes in the human MHC class III region near the junction with the clas  
interpart of mouse mammary tumor gene int-3.  
A:Reference number: A55562; MUID:95137587; PMID:7835890  
A:Accession: I61596  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EI  
A:Molecule type: DNA  
A:Residues: 1-404 <RES>  
A:Cross-references: GB:D28769; NID:G561657; PIDN:BAA05958.1; PID:G561659  
R:Nepper, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; J  
J. Biol. Chem. 267, 14998-15004, 1992  
A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation  
A:Reference number: A42879; MUID:92340547; PMID:1378943  
A:Accession: B42879  
A:Molecule type: mRNA  
A:Residues: 'G', 2-99, 'R', 101-404 <NEE>  
A:Cross-references: EMBL:M21211; NID:G190845; PIDN:AAA03574.1; PID:G190846  
A:Experimental source: lung  
A:Note: sequence extracted from NCBI backbone (NCBIP:109438)  
C:Comment: Advanced glycosylation end products are heterogeneous nonenzymatically glyco  
cellular function, thus contributing to tissue lesions in diabetes.  
C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide o  
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.  
C:Genetics:  
A:Gene: GDB:AGER  
A:Cross-references: GDB:306354; OMIM:600214  
A:Map position: 6p21.3-6p21.3  
A:Introns: 18/1; 53/3; 119/1; 140/3; 170/1; 231/1; 274/3; 322/1; 331/1; 373/2  
C:Function:  
A:Description: neuronal receptor for amphoterin, a DNA-binding protein involved in neu  
C:Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
C:Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-344/Domain: extracellular #status predicted <EXT>  
F:31-101/Domain: immunoglobulin homology <IM1>  
F:137-210/Domain: immunoglobulin homology <IM2>  
F:252-303/Domain: immunoglobulin homology <IM3>  
F:345-362/Domain: transmembrane #status predicted <TMW>  
F:363-404/Domain: intracellular #status predicted <INT>  
F:25,81/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:38-99,144-208,259-301/Disulfide bonds: #status predicted

Query Match 7.2% ; Score 147.5; DB 1; Length 404;

Best Local Similarity 22.6%; Pred. No. 0.0019;  
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;  
QY 15 EVIEGPQATVLKGSQARFNCTVSQW---WKLIIMWALSDMVVLSVRPMEPIITNDRTFS- 70  
Db 125 EIVDSASELTA--GVPNKVGTCTVSGSYPACTLSWHLG-----KPLVPNEKGVSV 173  
QY 71 ----QRYDGGNFT--SEMIHNVPSDSGNIR-----CSLQNSRLHGSAYITVQVMGELF 120  
Db 174 KEQTRRHPTGLFTLQSELM--VTPARGGDRPRTFSCFSFGLPRHRLARTAPIQREVM 230  
QY 121 IP-----SVNLVAENEP-----CEVTCPSHWTRLPDISWELGLLVSHSY 162  
Db 231 EPPLEEVOLV---EPEGAVAPGTVTLTCEVPAQPS-----PQIHWMD----- 274  
QY 163 YFVPEPSDLOSVALTPQSNGLTCTVATWKSLSKARKSATVNLTRCQDTGGGINI 222  
Db 275 -GVPLPLPPSPVLLIPEIGFODQGTYSVATHSHGPOESRAVSIIIE-PEEG----- 327  
QY 223 PGVLSLPSLPSLPTWGVGLAGTLLT-----PTCLTIRCCCRRCGCCGCC 275  
Db 328 -----PTAGSVGSGGLTALALGILGGLTAAALLIGVILWQR----- 366  
QY 276 CRCCFCRRKRGFRIOFKSKSKXT--NKETETSGNENSGYNSDEQKTTDTASLPKPS 333  
Db 367 -----QRRGERKAPENQEEERABLN----- 389  
QY 334 CESSDPEQRNNSCGPP 349  
Db 390 -QSEEPAGESSTGGP 404

RESULT 2  
151669  
tumor suppressor - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C;Accession: I51669  
R;Pierceall, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.  
Dev. Biol. 166, 654-665, 1994  
A;Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the  
A;Reference number: I51668; MUID:95113183; PMID:7613784  
A;Accession: I51669  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1427 <PIE>  
A;Cross-references: EMBL:U10986; NID:G606873; PIDN:AAA70168.1; PID:G606874  
C;Genetics:  
A;Gene: XDCCA

Query Match 6.5%; Score 133; DB 2; Length 1427;  
Best Local Similarity 25.7%; Pred. No. 0.088;  
Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;  
QY 17 IEFGQATVLKGSQARFNCTVSQW---KLIMWALSDMVVLSVRPMEPIITNDRTFSQRYD 74  
Db 243 LQPSNVVAIEGQDAVLECAVS-GYPTPIVMQGD-----EPVPPIR-----KYS 288  
QY 75 QGNGFTSEMIHNVPSDSGNIR---SLQNSRLHGSAYITVQVMGELFTIPSNLVVAENE 132  
Db 289 VLGG--SNLLISNVTDDAGAYTCVATYKNTSFSADLTVMVPPQLNHPANLYAESM 346  
QY 133 PCVETCL---PSHWTRLPDLSW-ELGLLVSHSYVFPPEPSDLOSVALTPQSNGT 187  
Db 347 DIFECVASKPS-----PVPKTKNGEVIPSDYFIQVDSNLR-----ILGLVKSDEGY 397  
QY 188 LTCVA-----TWKLSKARKSATVNLTVI-RCQDQGGGINIPGVLS 228  
Db 398 YQCTAENAGNIQTYAQLIIPDPAPVPSLSILSAPRDV-----VPLVSS 442

RESULT 3  
T24272

hypothetical protein T01B7.8 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T24272  
R;Sims, M.  
submitted to the EMBL Data Library, October 1995  
A;Reference number: Z19867  
A;Accession: T24272  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-164 <WIL>  
A;Cross-references: EMBL:Z66499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:T01B7.8  
A;Experimental source: clone T01B7  
C;Genetics:  
A;Gene: CESP:T01B7.8  
A;Map position: 2  
A;Introns: 20/3; 90/2  
Query Match 6.3%; Score 129.5; DB 2; Length 164;  
Best Local Similarity 32.1%; Pred. No. 0.015;  
Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;  
QY 153 LGLLVSHSYFYFPEPSDLOSVALTPQSNGLTCTVATWKSLSKARKSATVNLTVIRC 212  
Db 6 LAILLAIGTPIAV---SQQSAV-----LPVSSTELATVGTDVSTASTAIDTLGNSSRV 57  
QY 213 PQDTGGGINIPGVLSLPSLPSLPTWGVGLAGTLLTPTCLTIRCCCRRCGCCGCC 272  
Db 58 KRGGCGCGCGCG-----GCGCGCGCGG-----CGCCCRPRCCCRCCCTC 101  
QY 273 --NCCC-RCFCOR 283  
Db 102 CRTCCTRCCTCC 115  
RESULT 4  
156551  
neurotrophin - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 19-May-2000  
C;Accession: I56551  
R;Struyk, A.F.; Canoll, P.D.; Wolfgang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.  
J. Neurosci. 15, 2141-2156, 1995  
A;Title: Cloning of neurotrophin defines a new subfamily of differentially expressed nei  
A;Reference number: I56551; MUID:95198094; PMID:7891157  
A;Accession: I56551  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-344 <RES>  
A;Cross-references: EMBL:U16845; NID:G755184; PIDN:AAA67445.1; PID:G755185  
C;Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-term:  
Query Match 6.3%; Score 129.5; DB 2; Length 344;  
Best Local Similarity 24.2%; Pred. No. 0.034;  
Matches 64; Conservative 43; Mismatches 104; Indels 53; Gaps 14;  
QY 4 VIFLHGG-----SGNEVI-EGPQATVLKGSQARFNCTVSQWKLIMWALSDMVLS---- 55  
Db 21 LLFLVPTGVPRGDAFTFKAMDNTVRQGESATLCTIDNRTVRVAVLNRSTILYAGND 80  
QY 56 ---VRPEPIITNDRTFSQRYDQGNFTSEMITHNVPSDSGNIRCSLQ-----NSRLH 106  
Db 81 KWCLDPRVLLSN---TQTY-----SIEQNVVDVDEGPTTCVTDNHPKTSRVH 129  
QY 107 GSAVLTVQVMGELFIPSNLVVAENEPCEVTCLPSHWTRLPDISWELGLLVSHSYFYFVP 166  
Db 130 ----LIQVQSPKIVEISSDISINEGNISLTCTIATKPE-PTVTWR---HISPKAVGFVS 181  
QY 167 EPSDLOSVALTPQSNGLTCTVATWKSLSKARKSATVNLTVIRC-----QDTGGGIN 221  
Db 182 EDEYLE-----IQGITREQSGEYECSSAS-NDVAAVPRVRRVNVTVPPYISBAKGTGVPVG 236  
QY 222 IPGVL-----SSLPSLPSLPSLFTWGK 241

Db 237 QKGTQCEASAVPSAEFQ---WFK 257

RESULT 5

AS4100

N; tumor suppressor protein DCC precursor - human

N; Alternate names: colorectal cancer suppressor DCC

C; Species: Homo sapiens (man)

C; Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 05-Nov-1999

C; Accession: AS4100; A40098

R; Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B. Genes Dev. 8, 1174-1183, 1994

A; Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.

A; Reference number: A54100; MUID:95011532; PMID:7926722

A; Accession: A54100

A; Molecule type: mRNA

A; Residues: 1-1447 <HED>

A; Cross-references: EMBL:X761132; NID:G453209; PIDN:CAA53735.1; PID:G453210

R; Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton Science 247, 49-56, 1990

A; Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.

A; Reference number: A40098; MUID:90100559; PMID:2294591

A; Accession: A40098

A; Molecule type: mRNA

A; Residues: 1-750 <FEA>

A; Cross-references: GB:M32292; NID:G181492; PIDN:AAA35751.1; PID:G181493

C; Genetics:

A; Gene: GDB:DCC

A; Cross-references: GDB:119838; OMIM:120470

A; Map position: 18q21.1-18q21.1

C; Keywords: transmembrane protein; tumor suppressor

F; 1-25/Domain: signal sequence #status predicted <SIG>

F; 26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 6.2%; Score 128; DB 2; Length 1447;

Best Local Similarity 24.8%; Pred. No. 0.21; Indels 50; Gaps 10;

Matches 60; Conservative 33; Mismatches 99

Qy 10 SGSGNEV-----IEGPQNAVTLKGSQARFNTCTVSGW--KLIMWALEDM 51

Db 220 SRTGNEAEVRLSDPGIHLRQLYFLRPSNVVAIEGKDAVLECCVS-GYPPPSFTWLRGEE 278

Qy 52 VLVSVRPMELIITDFTSRDYDQGNFTSEMIHNVPSDSGNIRCSLONSRLHSGA 109

Db 279 VI-----Q-RSKYSLGG--SNLLISNVTDDSGMYTCVVYKNENISASA 323

Qy 110 YLTQVQNGELFIPSVNLVWAEPECEVTCLPSSHWTLPDISW-ELGLLVSHSSYFVPEP 168

Db 324 ELATVLPFPPFLNHPNLYAYESMDIEFECTVS-GKDPVTVMWKNGDVVPISDYFQVGG 382

Qy 169 SDIQSAVSIILATPQSNGLTLCVATWKSLSKARKSATNLTIVRCPODTGGGINIPGVLS 228

Db 383 SNLR-----ILGVVKSDEGFYQCAEAGNAGTSAQLIVPKPAIFSS-----VLPS 430

Qy 229 LP 230

Db 431 AP 432

RESULT 6

I48696

N; protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 2 - mouse

N; Alternate names: receptor-type tyrosine kinase

N; Contains: protein-tyrosine kinase nsk2 precursor, splice form 4

C; Species: Mus musculus (house mouse)

C; Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000

C; Accession: I48696; S60738

R; Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D. Oncogene 11, 281-290, 1995

A; Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase

A; Reference number: I48696; MUID:95349951; PMID:7624144

A; Accession: I48696

A; Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A; Residues: 1-881 <GAN1>

A; Cross-references: EMBL:X86445; NID:9929725; PIDN:CAA60166.1; PID:9929726

A; Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A; Molecule type: mRNA

A; Residues: 1-871 <GAN1>

A; Cross-references: EMBL:X86444; NID:9929723; PIDN:CAA60165.1; PID:9929724

A; Experimental source: splice form 2

A; Accession: S60738

A; Molecule type: DNA

A; Residues: 1-456 'A', 466-871 <GAN2>

A; Cross-references: EMBL:X86444; NID:9929723

A; Experimental source: splice form 4

C; Comment: For alternate splice forms see PIR:I48697.

C; Genetics:

A; Gene: nsk2

A; Cross-references: MGI:103308

C; Superfamily: mouse ror-related receptor; immunoglobulin homology; protein kinase hom

C; Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; ty

F; 1-21/Domain: signal sequence #status predicted <SIG>

F; 22-871/Product: protein-tyrosine kinase nsk2, splice form 2 #status predicted <MAT2>

F; 22-456 'A', 466-871/Product: protein-tyrosine kinase nsk2, splice form 4 #status pred

F; 42-101/Domain: immunoglobulin homology <IMM1>

F; 135-192/Domain: immunoglobulin homology <IMM2>

F; 226-284/Domain: immunoglobulin homology <IMM3>

F; 498-518/Domain: transmembrane #status predicted <TRM>

F; 575-865/Domain: protein kinase homology <KIN>

F; 583-591/Region: protein kinase ATP-binding motif

F; 222,462/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 6.2%; Score 127.5; DB 1; Length 871;

Best Local Similarity 21.7%; Pred. No. 0.13;

Matches 70; Conservative 41; Mismatches 125; Indels 87; Gaps 16;

Qy 15 EVIEGPQNAVTLKGSQARFNTCTVSGWK-LIMWALSVMVSVRPMELIITDFTSRDY 73

Db 122 KITRPPINVKIIEGLKAVLPCTTMGNPKPSVSMIKGD-----NALRENSRIAALE- 171

Qy 74 DQGNFTSEMIHNVPSDSGNIRCSLONSRLHSGA-----LTVQVWNGELFIPSVNLV 128

Db 172 -----SGSRKIHNVQEDAGQYRCVAKNSL--GTATSKLVKLEVEVLGRLRAPESHNV 223

Qy 129 AENPFCEVTC---LPSSHWTLPDISW-ELGLLVSHSSYFVPEPSDQSAVSIILATPQ 183

Db 224 TFGSFVTLRCTEIGIP-----VPTISWENGNVSSGSIQESVKDRVIDSRQLQFITKP- 277

Qy 184 SNGTLTCVAT-----WKSLSKARKSATNLTIVR-----CPDITGGGINIG- 225

Db 278 --GLYTICIANKHGEKFPSTAKAAATVSIASWKSQKDSQGYCAQYRGEGVLMQPGSKML 335

Qy 226 LSSLPISLGFSLP-----TWGKVLGLAGTLMLTPTCTLTITRCCCRRCGCCGN-- 273

Db 336 LVFLPTTSHRDPEDAQELLHTAWNEL-----KAVSPLCRPAEALLCVHLFLECSPG 388

Qy 274 -----CCCRCC-----PCCR 283

Db 389 VVPTMPICREYCLAVKELFCAK 411

RESULT 7

I48697

N; protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 1 - mouse

N; Alternate names: receptor-type tyrosine kinase

N; Contains: protein-tyrosine kinase nsk2 precursor, splice form 3

C; Species: Mus musculus (house mouse)

C; Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000

C; Accession: I48697; S60740

R; Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D. Oncogene 11, 281-290, 1995

A; Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase

A; Reference number: I48697; MUID:95349951; PMID:7624144

A; Accession: I48697

A; Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A; Residues: 1-881 <GAN1>

A; Cross-references: EMBL:X86445; NID:9929725; PIDN:CAA60166.1; PID:9929726

A:Experimental source: splice form 1  
 A:Accession: S60740  
 A:Molecule type: DNA  
 A:Residues: 1-456,'A',466-881 <GAN2>  
 A:Cross-references: EMBL:X86445; NID:g929725  
 A:Experimental source: splice form 3  
 C:Comment: For alternate splice forms see PIR:148696.  
 C:Genetics:  
 A:Gene: nsk2  
 A:Cross-references: MGI:103308  
 C:Superfamily: mouse ror-related receptor; immunoglobulin homology; protein kinase homology  
 C:Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; tyrosine  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-456,'A',466-881/Product: protein-tyrosine kinase nsk2, splice form 1 #status predicted <WAT>  
 F:42-101/Domain: immunoglobulin homology <IMM1>  
 F:42-456,'A',466-881/Product: protein-tyrosine kinase nsk2, splice form 3 #status predicted  
 F:135-192/Domain: immunoglobulin homology <IMM2>  
 F:226-284/Domain: immunoglobulin homology <IMM3>  
 F:498-518/Domain: transmembrane #status predicted <TRM>  
 F:575-865/Domain: protein kinase homology <KIN>  
 F:583-591/Region: protein kinase ATP-binding motif  
 F:222,462/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.2%; Score 127.5; DB 1; Length 881;  
 Best Local Similarity 21.7%; Pred. No. 0.13;  
 Matches 70; Conservative 41; Mismatches 125; Indels 87; Gaps 16;

QY 15 EVIEGPNATVLKGSQARFNCTVSGQWK-LIMWALS DMVLSVRPMEPIITNDRTSORY 73  
 DB 122 KTRTPINVKIIEGLKAVLPCTTMGNPKPSVNIKGD-----NAURENSRLAUE- 171  
 QY 74 DQGNFTSEMIHNVPDSGNIRCSLONSRLHGSAY-----LTQVMGELFIPSVNLV 128  
 DB 172 -----SGSLRIHNQKEDAGQYRCVAKNSL-GTAYSKLVKLEVEVLGRILRAPESHNV 223  
 QY 129 AENEPECEVTCL-PSHWTPLPDISWEL-GLLVSHSHSYFVPEPSDLSQSAVSIILATPQ 183  
 DB 224 TFGSFVTLRTEIGP-----VPTSIWENGNAVSGSIOESVKDRVIDSRILQIFTRP- 277  
 QY 184 SNTLTCTVATWKS LKARKSATVNLTVIR-----CPQDTGGGINTIPG-----V 225  
 DB 278 --GLYTCIATNKHGEKFTAKAAATVSIAEWSKSDSQGYCAQYREGVLMQGPGERKM 335  
 QY 226 LSLPLSLGRLP-----TWGKVLGLACTMLLTPTCTLTTRCCCRRCGCCN-- 273  
 DB 336 LVFLPTTSRDRPDAQELIHTANVEL-----KAVSLCPRAAEALLCYHLFLECSPG 388  
 QY 274 -----CCCRCC-----PCCR 283  
 DB 389 VVPTPMPICREYCLAVKLELFCAL 411

RESULT 8  
 T29757  
 protein UNC-89 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Dec-1999  
 C:Accession: T29757  
 R:Du, Z.; Lee, T.F.; Wilson, R.  
 submitted to the EMBL data Library, May 1997  
 A:Description: The sequence of C. elegans cosmid C09D1.  
 A:Reference number: T20679  
 A:Accession: T29757  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-6642 <DUZ>  
 A:Cross-references: EMBL:AF003131; PIDN:AA54132.1; GSPDB:GN00019; CESP:unc-89  
 A:Experimental source: strain Bristol N2; clone C09D1  
 C:Genetics:  
 A:Gene: CESP:unc-89  
 A:Map position: 1  
 A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1; 6/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match 6.1%; Score 125.5; DB 2; Length 6642;  
 Best Local Similarity 23.9%; Pred. No. 1.6;  
 Matches 52; Conservative 39; Mismatches 92; Indels 35; Gaps 8;

QY 16 VIEGPNATVLKGSQARFNCTVSGQW--KLIMWALS DMVLSVRPMEPIITNDRTSORY 73  
 DB 2077 VVDGPKSVTIKETETAEFKATIS-GFPAPTAKTINEKIVEESRTTITIKEDVYT----- 2131  
 QY 74 DQGNFTSEMIHNVPDSGNIRCSLONSRLHGSAYLTQVMGELFIPSVNLV 128  
 DB 2132 -----LKISNAKIEQTGVKVTAAQNSAGQDSKQADLKVEPNVAKAFKSKSLTDKVA 2182  
 QY 129 AENEPECEVTCL-PSHWTPLPDISWEL-GLLVSHSHSYFVPEPSDLSQSAVSIILATPQ 184  
 DB 2183 DEGEFLRNWLELDGSPSGT---EVSMLLNGQPTKSTQVVDHGGTYHTVIAEAKPEM 2239  
 QY 185 NGTLTCTVATWKS LKARKSATVNLTV-----VIRCPQD 215  
 DB 2240 SGTLTAKAKNAAGECETS AKVTVNGGNKKPEFVQAPQN 2277

RESULT 9  
 JC4593  
 protein-tyrosine kinase-related receptor PTK7 precursor - human  
 N:Alternate names: receptor protein tyrosine kinase-like protein (RPTK)  
 C:Species: Homo sapiens (man)  
 C:Date: 16-Apr-1996 #sequence\_revision 24-May-1996 #text\_change 24-Sep-1999  
 C:Accession: JC4593  
 R:Park, S.K.; Lee, H.S.; Lee, S.T.  
 J. Biochem. 119, 235-239, 1996  
 A:Title: Characterization of the human full-length PTK7 cDNA encoding a receptor prote  
 A:Reference number: JC4593; MUID:97037064; PMID:8882711  
 A:Accession: JC4593  
 A:Molecule type: mRNA  
 A:Residues: 1-1070 <PAR>  
 A:Cross-references: GB:U40271; NID:gl322231; PIDN:AAC50484.1; PID:gl322232  
 C:Comment: This protein is a member of receptor protein tyrosine kinase family, but pr  
 C:Genetics:  
 A:Gene: GDB:PTK7  
 A:Cross-references: GDB:134760; OMIM:601890  
 A:Map position: 6p21.1-6p21.2  
 C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homo  
 C:Keywords: cell adhesion; extracellular protein; glycoprotein; phosphotransferase; tr  
 F:1-30/Domain: signal sequence #status predicted <SIG>  
 F:31-1070/Product: protein-tyrosine kinase 7 #status predicted <NAT>  
 F:31-703/Domain: extracellular #status predicted <EXC>  
 F:704-725/Domain: transmembrane #status predicted <TM>  
 F:726-1070/Domain: intracellular #status predicted <INC>  
 F:794-1065/Domain: protein kinase homology <KIN>  
 F:116,175,184,214,268,283,405,463,567,646/Binding site: carbohydrate (Asn) (covalent)

Query Match 6.0%; Score 124; DB 2; Length 1070;  
 Best Local Similarity 25.3%; Pred. No. 0.3;  
 Matches 58; Conservative 32; Mismatches 103; Indels 36; Gaps 10;

QY 16 VIEGPNATVLKGSQARFNCTVSGQWKLIMWALS DMVLSVRPMEPIITNDRTSORY 74  
 DB 227 VVLAPQDVVARYEAMFHCQFSQPPSLQMLFEDETPITNRSRPPHLRATVFA----- 282  
 QY 75 QGDNFTSEMIHNVPDSGNIRCSLONSRLHGSAYLTQVMGELFIPSVNLV 129  
 DB 283 -----NGSLLTQVPRNAGIYRCIGQGRGPPILATLHAEIEDMPLEFVRVFTAGS 337  
 QY 130 ENEPECEVTCLPSHWTPLDISWE-LGL-LVSHSHSYFVPEPSDLSQSAVSIILATPQSN-G 186  
 DB 338 EE--RVTCLEPKGLPEPSVWHEAGVRLPTHGRVY-----QKGHELVLIAESDAG 387  
 QY 187 TLTCVATWKS LKARKSATVNLTV-----TRCPDQGGGINIPGVLSL 229  
 DB 388 VYTCHAA--NLAGORRQDVNITVATVPSWLKPKQDSQLBEKPGYLDCL 434

## RESULT 10

TI5651  
 Hypothetical protein C27A2.5 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: TI5651  
 R:Nhan, M.  
 submitted to the EMBL Data Library, May 1996  
 A:Description: The sequence of C. elegans cosmid C27A2.  
 A:Reference number: Z18382  
 A:Accession: TI5651  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-188 <NHA>  
 A:Cross-references: EMBL:U58760; NID:gl330384; PID:gl330389; PIDN:AAB00710.1; GSPDB:GN00  
 A:Experimental source: strain Bristol N2; clone C27A2  
 C:Genetics:  
 A:Gene: CESP:C27A2.5  
 A:Map position: 2  
 A:Introns: 19/3; 91/2

Query Match 5.9%; Score 122.5; DB 2; Length 188;  
 Best Local Similarity 61.3%; Pred. No. 0.059; 8; Indels 3; Gaps 2;  
 Matches 19; Conservative 1; Mismatches 102; Indels 74; Gaps 18;

QY 256 CTTTIRCCCRRCRCGC--NCOC-RCCFCRCR 283

DB 86 CCRPKCCCCRCCTCCCTCCCTCCCTCC 116

## RESULT 11

JC4025  
 opioioid-binding cell adhesion protein - human  
 C:Species: Homo sapiens (man)  
 C>Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 17-Mar-2000  
 C:Accession: JC4025  
 R:Shark, K.B.; Lee, N.M.  
 Gene 155, 213-317, 1995  
 A:Title: Cloning, sequencing and localization to chromosome 11 of a cDNA encoding a huma  
 A:Reference number: JC4025; MUID:95237612; PMID:7721093  
 A:Accession: JC4025  
 A:Molecule type: mRNA  
 A:Residues: 1-345 <SHA>  
 A:Cross-references: GB:I34774; NID:9514373; PIDN:AAA36387.1; PID:g514374  
 A:Experimental source: brain  
 C:Comment: This protein binds opioioid alkaloids in the presence of acidic lipids, exhibi  
 C:Genetics:  
 A:Gene: GDB:OPCML; OBCAM; OPCM  
 A:Cross-references: GDB:251677; OMIM:600632  
 A:Map position: llpter-llqter  
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 5.8%; Score 120; DB 2; Length 345;  
 Best Local Similarity 25.0%; Pred. No. 0.17;  
 Matches 71; Conservative 37; Mismatches 102; Indels 74; Gaps 18;

QY 4 VIFLHSGS----SGNEVI-EGPQNAVTLKGSQARFNCTVSGQWKLIMWLSDMVVL----NSRL 105

DB 21 LLFLVPTGVPVRSGDATFPKAMDNVTVRQGESATLCTIDDRVTRVAV-LNRSTILYAGN 79

QY 55 ---SVRPMEPIITNDRTSORYDQGNFTSEMIIHNVEPSDSGNIRCSLQ-----NSRL 105

DB 80 DKWSIDPRVILVN---TPTQY-----SIMIQNVVDYDEGPTCSVQTDNHPKTSRV 128

QY 106 HGSAYITVQVWGELFIPSNLVVAENPEVTCV-----PSHWRTRLPDISWELGLLVSHSS 161

DB 129 H-----LIVQVPPQIMNISDITVNEGSSVTLCLATGRPE-----PTVTR-----HLS 173

QY 162 YY----FVPEPSDLQSAVSLTALTPQSNGLTLCVATWKSILKARKSATVNLTV-----IRC 212

DB 174 VKEGQGFVSEDEYLE----ISDIKRDQSGEYECAL-NDVAAPDVRVKVITVNPYPYISK 228

QY 213 PQDTGGGINTPGVL-----SSLPSLGFSLPTWKG-----VGLGLAG 248

DB 229 AKNTGVSQKGLSCASAVPMABEQ---WPKETRLATGLDG 269

## RESULT 12

S03199  
 opioioid-binding protein OPCAM precursor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 17-Mar-2000  
 C:Accession: S03199  
 R:Schofield, P.R.; McFarland, K.C.; Hayflick, J.S.; Wilcox, J.N.; Cho, T.M.; Roy, S.;  
 EMBO J. 8, 489-495, 1989  
 A:Title: Molecular characterization of a new immunoglobulin superfamily protein with p  
 A:Reference number: S03199; MUID:89251576; PMID:2721489  
 A:Accession: S03199  
 A:Molecule type: mRNA  
 A:Residues: 1-345 <SCH>  
 A:Cross-references: EMBL:X12672; NID:g5985; PIDN:CAA31192.1; PID:g5986  
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-term  
 C:Keywords: transmembrane protein  
 F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:28-345/Product: opioioid-binding protein OPCAM #status predicted <MAT>

Query Match 5.8%; Score 119; DB 2; Length 345;  
 Best Local Similarity 25.1%; Pred. No. 0.2;  
 Matches 70; Conservative 36; Mismatches 103; Indels 70; Gaps 17;

QY 4 VIFLHSGS----SGNEVI-EGPQNAVTLKGSQARFNCTVSGQWKLIMWLSDMVVL----NSRL 105

DB 21 LLFLVPTGVPVRSGDATFPKAMDNVTVRQGESATLCTIDDRVTRVAV-LNRSTILYAGN 79

QY 55 ---SVRPMEPIITNDRTSORYDQGNFTSEMIIHNVEPSDSGNIRCSLQ-----NSRL 105

DB 80 DKWSIDPRVILVN---TPTQY-----SIMIQNVVDYDEGPTCSVQTDNHPKTSRV 128

QY 106 HGSAYITVQVWGELFIPSNLVVAENPEVTCV-----PSHWRTRLPDISWELGLLVSHSS 161

DB 129 H-----LIVQVPPQIMNISDITVNEGSSVTLCLATGRPE-----PTVTR-----HLS 173

QY 162 YY----FVPEPSDLQSAVSLTALTPQSNGLTLCVATWKSILKARKSATVNLTV-----IRC 212

DB 174 VKEGQGFVSEDEYLE----ISDIKRDQSGEYECAL-NDVAAPDVRVKVITVNPYPYISK 228

QY 213 PQDTGGGINTPGVL-----SSLPSLGFSLPTWKGVLGLA 247

DB 229 AKNTGVSQKGLSCASAVPMABEQ---WPKEDTRLA 264

## RESULT 13

JC1239  
 opioioid-binding protein (clones SG8 and SG13) - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-May-2000  
 C:Accession: JC1239  
 R:Lippman, D.A.; Lee, N.M.; Loh, H.H.  
 Gene 117, 249-254, 1992  
 A:Title: Opioioid-binding cell adhesion molecule (OBCAM)-related clones from a rat brain  
 A:Reference number: JC1238; MUID:92347701; PMID:1339369  
 A:Accession: JC1239  
 A:Molecule type: mRNA  
 A:Residues: 1-345 <LIP>  
 A:Cross-references: GB:M88710; NID:g203247; PIDN:AAA40859.1; PID:g203248; GB:M88711; N  
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-term  
 C:Keywords: transmembrane protein

Query Match 5.7%; Score 118; DB 2; Length 345;  
 Best Local Similarity 24.7%; Pred. No. 0.24;  
 Matches 71; Conservative 38; Mismatches 104; Indels 74; Gaps 18;

QY 4 VIFLHSGS----SGNEVI-EGPQNAVTLKGSQARFNCTVSGQWKLIMWLSDMVVL----NSRL 105

DB 21 LLFLVPTGVPVRSGDATFPKAMDNVTVRQGESATLCTIDDRVTRVAV-LNRSTILYAGN 79